Philosophical Talent

Empirical investigations into philosophical features of adolescents' discourse

Thecla Rondhuis

Philosophical Talent

Empirical investigations into philosophical features of adolescents' discourse

Filosofisch Talent

Empirisch onderzoek naar filosofische kenmerken in gedachtenwisselingen van jongeren
(met een samenvatting in het Nederlands)

Proefschrift ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de Rector Magnificus, Prof.dr. W.H. Gispen,
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen op 30 september 2005 des middags te 14.30 uur
door Niesje Thecla Wijnanda Rondhuis, geboren op 30 januari 1950,
te Hilversum.

Promotor:
Prof.dr. P.G. Heymans, Faculteit Sociale Wetenschappen, Afdeling Ontwikkelingspsychologie, Universiteit Utrecht

Copromotor:
Dr. K.L. van der Leeuw, Faculteit der Geesteswetenschappen, Afdeling Wijsbegeerte, Universiteit van Amsterdam

Financiële ondersteuning: Universiteit Utrecht, Faculteit Sociale Wetenschappen

ISBN nummer: 90-393-4033-1

Foto omslag: Jean-Pierre Jans

Vormgeving: Jos Rietstap Vormgeving, Schiedam

Druk: Veenman Drukkers, Rotterdam

Contents

1 Philosophical talent — 1
1 Conceptualisation of talent — 3
2 What does philosophical talent mean? — 5
3 Why look for philosophical talent among youngsters? — 5
4 Empirical identification of philosophical talent: the design of this research project — 7

2 Characterisations of philosopher's activity — 9
1 Philosophy, philosophising, and being wise — 9
2 Characterisations of philosopher's activity in the 20th century — 12
3 Main features of philosopher's activity — 14
4 Wisdom and psychological approaches of the 20th century — 16
5 Main features of philosopher's activity and wisdom as thinking patterns — 21
6 Summary and discussion — 23

3 Thinking pattern: Bridging philosophy and psychology — 25
1 Thinking themes and patterns of children: samples — 26
2 Evaluative approaches of children's answer finding procedures — 40
3 Approaching children's answer finding procedures beyond conventional norms — 45
4 Measuring thinking patterns — 49
5 Conclusion — 51

4 Assessing Philosophical Quality — 53
1 PQ Indicators: a conceptual frame-work — 53
2 The Tetralogue: a standardised procedure for assigning PQ indicators — 62
3 Investigations into the psychometric qualities of the Tetralogue — 66
4 Results concerning objectivity, reliability, and validity of the measurements — 69
5 Conclusion — 97

5 Similarities and differences between philosophical themes — 99
1 Different themes of the tetralogues — 99
2 Academic philosophy and tetralogue philosophy — 104
3 Similarities in the character of inquiry — 106
4 Variations in tetrlogue characteristics in function of philosophical themes — 108
5 Conclusion — 110

6 Participant's characteristics in relation to philosophical quality — 111
1 Validation of the tetralogue — 111
2 Theoretically expected relationships between participant's characteristics and philosophical quality — 113
3 Methods to measure and analyse — 121
4 Results related to construct validation — 129

5 Exploring relations between philosophical quality and participant's characteristics — 131
6 Discussion about potential effects of age, gender, irregular life course, and threshold value of convergent thinking — 132
7 Summary and conclusion — 137

7 Chair interventions and general didactics of tetralogues — 139
1 Philosophical discussions — 139
2 Didactics of tetralogues — 141
3 Differences between tetralogues related to chairs — 142
4 Differences between tetralogues related to numbers of interventions — 143
5 Conclusions — 143

8 Two-year follow up of PQ — 145
1 Observing children in their transition from primary to secondary school — 145
2 Selection follow-up participants — 147
3 Nine successive tetralogues — 147
4 Explorations of the longitudinal data — 151
5 Conclusion and discussion — 156

9 Philosophical quality in perspective — 159
1 Philosophical quality, pq and PQ — 159
2 Philosophical quality in the perspective of philosophy — 162
3 Philosophical quality related to rationality — 163
4 Philosophical quality vis-à-vis doing philosophy with children — 166
5 Philosophical quality in the perspective of empirical studies — 168
6 Philosophical quality as talent — 171
7 Applicability of philosophical quality — 172

References — 177

Appendices — 183

Summary — 197
Samenvatting — 203

Cooperating schools and institutes — 209

Acknowledgement — 211

Curriculum vitae — 213

1 Philosophical talent

Thinking patterns defined by well-known philosophers can be often recognised in trains of thought uttered by children and youngsters. For example, a ten year-old girl once was asked by her mother to clean her room. She was heard to be rumbling around behind the closed door of her room, then suddenly fell silent and answered after a while: 'Does a mess exist when nobody sees it?' Although the child was not familiar with great movements within epistemology, her question seems to refer to the discussion between rationalism and empiricism. Other examples will be presented in Chapter 3. How certain some children and youngsters demonstrate their sense of inquiry, for example into the essence of a horse, resembles Aristotle's careful observation of the sex of a hyena (Aristotle, 2003). Do these youngsters and philosophers share some intellectual turmoil, wonderment, persistent and consistent thinking patterns, and inclination to search for ambiguities in real life experiences? This topic will be examined in this study.

Recorded philosophical discussions with twelve year olds are simultaneously funny and serious because youngsters may readily accept unreal and ethereal conclusions of thinking trials. Judgements classified as *funny* or *serious* are often based on ambiguities, vagueness and uncertainty of utterances, and situations viewed from different angles. Images seen from one angle may conflict with atmosphere, expectations or context of the other. For example, the comments of the 12 year old boy in Chapter XII of *The Little Prince* (Saint-Exupéry, 1980) in which a character is reported to drink many bottles of wine: 'If a drunk keeps on drinking, he will become old' (Rondhuis, 2001). Although some trains of thought construct circular or contradictory arguments, they also uncover hidden standards. Another 12 year-old suggested that things can be made bigger by placing them under a microscope and subsequently inventing a formula to assess the expanded version as compared to the real (see Chapter 3: Can countries on earth change?). Most adults find such expressions humorous. On second thought however, it turns out that only different standards are employed to value the proportions of things. Is size measured by eye, by touch, or both? There must be something that makes people sensitive to ambiguity, to interpret objects, situations and events in life in multiple senses, to view them from other potential angles. Perhaps it is this particular ability that adds philosophical quality to discussions and offers the impression of being a talent.

Inspiration to this inquiry

Numerous philosophical discussions with 11 to 12 year old children and with youngsters from secondary schools inspired me to investigate the philosophical quality of their thinking patterns. These thinking patterns are captivating and often reflect classical philosophical ideas. At the same time, one may recognise in these patterns, trains of thought performed in one's own youth. Frequent observations of their questioning and their basic exploratory behaviour led to the question of why are some children and youngsters sensitive to philosophical questions and others not. There must be something in the mind that generates wonderment and

evokes the autonomous production of authentic thoughts that are philosophically qualified. Perhaps it is the same fascination expressed by Daniel Dennett (1993) when questioning his audience: 'Why are you actually so surprised that a scientific career is nothing more than an enlargement of children's questions that are not yet answered?' This is something that might be formulated as a talent.

Autobiographic inspiration

Many thinking patterns expressed by children and youngsters today also reflect my own trains of thought when I was a primary school child. Growing up as a Roman Catholic I had to go to confession, communicating all my sins to a priest. Subsequently, I had to promise never to sin again and was forced to conduct the penitential imposed upon me by the priest consisting of some prayers. Before entering the confession box, I was supposed to review my sins over the previous month. As my life was not very spectacular, neither were my sins. Reflecting on this for a while, I wondered especially about two things. The first was my repetitive sin of disobedience to my parents. Whom should my mother obey? That must have been my father. But whom should my father obey? His chief? Or the boss of his chief? But then, whom should the world's chief (perceived as Eisenhower at that time) obey? I asked my mother and she promised me to think about it. Maybe it was God. But what if world's leader did not believe in God, or what if people did not believe in the obligation of obeying anybody? There were two options for me: either all the world leaders (past, present and future) must believe in the same God or that there was an endless regression neglecting any God. The option that the world's leader really was at the top of the hierarchy did not enter my mind. The first option could be proven invalid, as I knew Caesar did not believe in the same God as I had grown up with. Therefore, the second option must have been valid, igniting my disbelief from then on.

I remember very well my doubts, the inaccessibility of answers and my teachers' judgement that such questions were impolite. My second wonderment dealt with the discrepancy between promising and doing, experiencing a needless link between these faculties like all children. For instance, children played with marbles at school, but not always in an honest way. All children, I thought, communicated this unfair behaviour with the priest and subsequently promised to be honest in the future, although this rarely happened. In my childish opinion, a necessary relationship between promise and action must exist because of the very meaning of a promise. The alternative possibility I could imagine was that the confessional was a sham.

Countless examples of similar children's trains of thought are reminiscent of Descartes' experiment of doubting (1979); Nietzsche arguing God's death (1999); Kant's effort to overcome the controversy between rationalism and empiricism (2004); or Wittgenstein's listing of the trivials of daily life (1973). These philosophically qualified thinking patterns and questions are challenging through their openness, their playful approach of reality, full of ambiguities and game like reasoning, sometimes with a rigid logic, and always bridging between real life experiences to domains of abstractions and *vice versa*. The search for philosophical

quality that may be interpreted as a talent will begin by identifying the concept of *talent* (Section 1.1). Next focus will be on the meaning of *philosophical talent* (Section 1.2). As this quest for philosophically qualified thinking patterns is done with young people, considerations for selection of this group are given in Section 1.3. The existence of a philosophical talent and the status of research outcomes are discussed. Finally, an overview of the research project is presented (Section 1.4).

1 Conceptualisation of talent

It is easy to reach a consensus about to whom talent can be ascribed. Aristotle, Leonardo da Vinci, Shakespeare and Beethoven share a form of successful performance that differs from most conventional achievers, irrespective the object of performance (academic, social, etc.). The Oxford Dictionary describes three meanings of *talent*: 1) denomination of weight, the value of it or treasure, 2) inclination, disposition, 3) mental endowment. Although the last definition will be used here, this description is still too general, vague and indistinct. At the same time, the similarity between Aristotle and Beethoven still escapes the ontological source of the concept *talent*. An Internet search of the question 'What is a talent?' generates 273 thousand pages (probably much more at the very moment of reading this), varying from carrier coaching to recognising talent in canoe slalom. Even the question 'What is philosophical talent?' results in 878 pages with references. Talent may be approached according to different discourses: in everyday use; empirically as in many psychological studies; or metaphorically, as in the bible (in different meanings). The concept of *talent* will be approached here in a Socratic way: analysing the use of the concept in concrete examples and studies about this topic, seeking for similarities and dissimilarities between people who are considered to be talented at different levels, and between objects in different empirical studies.

Concrete examples

Can talents refer to every indiscriminate quality? Is Bill Gates talented because he is successful in making money? Or was Napoleon talented in conquering so many lands and making strategic choices? Is it possible being talented in tiling, in performing magic tricks, in kingship, or in inventing new languages (e.g., Esperanto)? Is every performance potentially qualified as talent? How does the potential develop into achievement? Has successful performance to be developed? In this respect, successful performance has something in common with the object of education. According to Peters (1980), the object of education is not just a certain activity. It is related to criteria that have to be satisfied in order to assign the quality of being successful. Criteria emerge when resulting achievements are differentiated between those that can be related to potential talents and those that cannot. A successful performance related to a talent is valuable in a special sense. Some qualities are valued more than others, being more appropriate as a talent referent. Several academic achievements are valued to similar degrees, while they do not equally carry the potential of a talent. Achievements in mathematics,

literature and arts are possibly linked to talent, while the quality of talent is open to question in case of medical or economical achievements. Qualities that are sensitive to being classified as a talent are chiefly restricted to those with clear cognitive meanings such as high analytics, music, moral excellence and mathematics. In the same volume (Peters, 1980), Hirst distinguishes a number of fundamental cognitive categories: mathematics, physics, humanities, history, religion, literature, arts and philosophy, each with its own terms, usage, and justification. If talents refer to such specialised concept systems, a philosophical talent may also be formulated.

Talent as objects in empirical studies

As in many psychological studies, talent is approached empirically. Some empirical attempts are made to comprehend talents in music, in mathematics, in general intelligence, or in successful personality traits. In Developing Talent across the Lifespan (Van Lieshout & Heymans, 2000), the authors emphasise that talent is no longer viewed as just a matter of rather stable individual differences in potential or performance. Instead, developmental changes in talents are currently examined in close relation to changing contextual support, changing constraints, and changing tasks. Talent is supposed to result from the acquisition of a sequence of skills, facilitated by changes in the individual's environment. The authors try to find which aptitudes and environmental features are involved in the acquisition of these skills and how these fit with the individuals' development. A talent may be an emergent construct, developing dynamically and chaotically with lots of varying parameters and sensitive starting conditions. The question for a talent's source refers to another empirical issue too. Can talent belong to the genetic heritage, and considered in a biological discourse? Or is it just the acquisition of a sequence of skills? In this respect, Matthew 25: 14-30 relates that 'talents' lent by the Lord appear to lead to success if, and only if, they are exploited in cooperation with the environment.

Reviewing the aspects found, it must be stressed that *talent* concerns successful performance with sensitive starting conditions, an ability that must be exploited in cooperation with the environment, and a valuable quality that is rooted in rational beliefs. However, one aspect needs to be emphasised in addition. As described in the introduction, it was the experience of philosophical dialogues with children and youngsters that made me think of apparent philosophical talent. Something in the utterances of youngsters was recognised as promising in a philosophical sense. Obviously, talent suggests a promise. In contrast to the elderly, children may be qualified as potentially talented. Some children will be identified as talented in whatever achievement, even without corresponding success. Successful performance of a talented individual does not need to be immediately realised. It may be realised in the future, with the interplay between experience and environment. Talented children and celebrities share a potential recognised in their utterances of promising mental endowment, but may differ in having demonstrated the corresponding successful performance.

2 What does philosophical talent mean?

The identification of potential philosophical talent will be the object of this study. What does philosophical talent mean? Philosophical talent may be attributed to Descartes or Popper, but does the title of philosopher guarantee a philosophical talent? Any attempt of describing a philosophical feature is probably doomed to fail because it will exclude some well-known philosophically qualified people. On the other hand, many such descriptions are elusive. Universal qualities like *domain transcending thinking*, or *the ability to view things from different perspectives*, evoke questions such as 'which domain must be transcended and which kinds of perspectives are more or less appropriate?' At the same time, well-known thinkers without the title of philosopher, like Darwin or Piaget need not be excluded from being philosophically talented.

Why is philosophical talent not simply reasoning? As Russell (1981) pointed out, philosophy is the no-man's-land between theology and sciences (i.e., all definitive knowledge), open to all attacks from both sides. The examples in the above sections show philosophically qualified thinking patterns that model the objective of this research. They imply some similarities with classic philosophical ideas and attitudes that are open to questions originating from the rational as well as those from passion. At the same time, they astonish and impart a degree of spiritual and intellectual uncertainty for the researcher. Emphasising some introductory remarks, the investigated philosophical talent will be characterised by general descriptions like the autonomous production of thoughts resulting from wondering, together with inquisitive, tentative and sensitive searching. Philosophising brings presuppositions up for discussion, concepts are analysed, and ambiguities, vagueness and uncertainty are uncovered. Exponents of philosophically qualified thinking patterns show a vulnerable attitude, are happy to disagree with themselves, avoid certainties or dogmas, and will defer or suspend generalised judgements. Chapter 2 reviews historical approaches to philosophically qualified thinking patterns and similarities between several philosophical performances, for example, academic philosophy and philosophising as mental activity.

3 Why look for philosophical talent among youngsters?

Philosophical talent involves a promise to perform philosophically qualified thinking patterns and must be detected in domains of expectations with respect to future performance. Searching for these patterns in promising young people will trigger a philosophically risqué overture. Observing the thinking patterns of youngsters and elaborating on their trains of thought inspired me to undertake this research. The selection of the research group must be adequate in its prototypical sense, but also with respect to its general power of expressing a philosophical talent. Besides the advantages of investigating youngsters' philosophical thinking patterns illustrated through relating anecdotes, their limitations and the evident differences from adult thinking patterns must be explained. Children and youngsters are a very suitable group to communicate with during philosophising exercises because they have less experience in classifying ordinary things than adults and can therefore be

genuinely surprised by seemingly trivial events. Provided with fewer acquired facts and values, knowledge and opinions, it is easy for them to investigate these events in an unprejudiced way, a primary condition for the intellectual exercise embodied in philosophical dialogues. Many things in the world are new to them and classic opinions have not yet taken root. The borderline between familiar and absurd things is still vague for them. Although children are predominantly practical thinkers and although they usually do not like nebulous arguments, they are capable of fantasising incredibly although by no means blindly. A clear-cut distinction between thinking and action, between considering and experiencing does not exist for them.

Part of a child's approach to events is the intention to question, the number of questions, and the acceptance of all kinds of answers. The eagerness to enter into open searches is illustrated by young children approaching their parents with questions like: 'Is the car behind us also going to grandpa's and grandma's?' Adults would think: 'Of course not.' But they cannot be sure nor can they prove the converse. Rather, adults would never ask or consider such questions. Children put their questions forward in a trial-like manner, often spontaneously, sometimes as a result of spontaneous situations, or posed after long periods of consideration. 'Why do you appreciate monotonous handwriting yet detest a monotonous landscape?' 'Is it possible to think of things that do not exist?' Open questions are invitations to search for meanings rather than to seek answers. Philosophically qualified thinking patterns are the result of an inquisitive procedure that must follow questioning. Adults usually interpret questions within the limits of knowledge systems, preferring clear-cut answers. 'Mama, why are you a girl?' 'I am a girl because I was born with a double X-chromosome, from which a human body develops with the constitution of a girl.' Of course, such an answer is not the one the child is looking for, because definitive answers kill a child's ambition to investigate further.

Another example of the difference between adults and youngsters is given by the report of an eleven year old girl. 'We were with friends and I told them about philosophising. Then we were engaged in a philosophical discussion. But adults don't understand anything about it. For example, when I asked: 'How do whales feel themselves?' they answered: 'you have to ask Greenpeace.' I asked: 'What feelings do flowers have? Is a flower able to think and see?' Then they replied: 'You'd better ask a scientist.' Isn't that foolish? Now, my Grandpa is the only one to philosophise with.' Evoking philosophically qualified thinking patterns with children involves the development of an open inquiry. Children are testing their thoughts about people, objects and events against their experience, against the thoughts of others, or other potential thoughts, and against generally accepted opinions. Clichés are dropped because they are not rooted and are replaced by autonomously performed thinking patterns. Undertaking philosophical discourses with children involves challenging thinking patterns in the absence of definite answers that prevent the discussion from becoming just another teaching method to arrive at a definitive answer.

Notwithstanding the attractiveness of philosophising with children in this study, potential divergences in thinking patterns between the young and adults have to

be reviewed in order to estimate the adequacy and the limits of exploring adolescent thinking patterns. Children may differ from adults in their thinking procedures and in thinking-results because of imperfections in their knowledge according to adult standards and in the number of events experienced. In terms of the strength of verbal access to ideas, contrasts between adolescents and adults may be differences of degree. Thinking patterns of children, their creativity and curiosity are certainly not inferior to those of an adult. Neither are the philosophical issues discussed. The mysteries of daily life are common issues for both age categories. Differences that appear often concern socially controlled elements and the degree of incorporating formal thinking patterns into expressions of ideas. Pure, independent thinking seldom occurs. It is hard to escape from decorum. In exceptional cases, this may paralyse or hide genuine thinking. The hypothesis is that factors like decorum and socially controlled elements increase during adulthood because the young are less hampered by cultural baggage. It is easier for adolescents to think autonomously, whereas adults have to cast off their culturally imposed yokes in order to engage in genuine thinking.

This study profits from the factors mentioned above in its identification of philosophically qualified thinking patterns. On the one hand thinking patterns of youngsters are untouched by learned reaction patterns, knowledge and prejudices. Youngsters are genuinely surprised about daily events; they question, have unusual thoughts and like to be involved in limitless inquiry. Moreover, they are not yet subject to the social restrictions and cliché-ridden approaches to reality often observed in adults. Adolescents in the age range 10 to 20 years have developed a relatively broad spectrum of oral expressions that facilitate philosophical discussions. All the above factors make the ten to twenty year olds an adequate group for an investigation of philosophical talent.

4 *An empirical study into philosophical quality and talent*

For much of my professional career I have enjoyed observing youngsters as they make their first excursions into the philosophical realm, as they become entrained by their curiosity and exploite mental experiences through the performance of inquisitive trains of thought. To this end, I collected and analysed observable thinking patterns and attempted to predict how participants would perform in future dialogues. For children and youngsters, performing philosophically qualified thinking patterns is an exciting experiment. Something is under construction: physical, mental or spiritual moves are in the process of being linked up in a constructive way. Some youngsters are apparently more apt than others: they generate thinking patterns differing in philosophical quality. It might be that youngsters differ in their mental endowment for philosophising, or in their philosophical talent.

The research project 'Philosophical Talent' attempts to answer the following questions:
1. Which features signify philosophical quality?
2. How can philosophical quality be assessed?

3. Can the philosophical quality of an individual's systematic thinking patterns be interpreted as a measure of philosophical talent?

An overview of several theoretical and empirical investigations undertaken as part of this research project is presented below.

First, philosophical characteristics are analysed theoretically, against the background of historical developments in philosophy, in addition to descriptions about the mental activity of philosophising by modern philosophers. This leads to the formulation of the main *features of philosophy* (Chapter 2). Charged with these theoretical findings, *real life expressions of youngsters* are explored (Chapter 3). They are observed precisely with respect to their content, as well as to their procedures. Observable utterances by youngsters are evaluated in order to recognise philosophical features in these expressions.

In Chapter 4, a series of clear *observable indicators* is developed to identify philosophically qualified thinking patterns. Subsequently, an *instrument* is created through which philosophically significant indicators can be assessed. This instrument is called *tetralogue*; it comprises a philosophical discussion by four youngsters in which they exchange their trains of thought in response to a philosophical topic, and which is overseen by a chairperson. *Data* are *collected* for about 100 discussions with close to 300 participants. Each discussion is undertaken by a small number of participants, preferably four, with registered characteristics concerning age, educational level, irregular life course and gender. Recorded discussions are transcribed and scored. Subsequently, the tetralogue is tested with respect to its *objectivity of scoring, reliability and validity*. Philosophical qualities are measured on two performance levels: individual and group. A *numerical formula* is created to describe the complexity of theoretical considerations of philosophical quality by means of the observed indicators. A second formula is developed to express the philosophical power of the group performance that reflects more than individual contributions.

In subsequent chapters (chapters 5 to 8), empirical manifestations of philosophical qualities of individual and group performances are related to other attributes of individuals and groups. This set of empirically obtained relationships is compared with the assumed relationships from the *nomological network* linking the construct 'philosophical quality' to several characteristics of the situation, participants, and discussion group in order to further corroborate the construct validity of the measurement instrument. After having checked the influence of different *philosophical topics* (Chapter 5), the focus is directed towards characteristics of participants and group performances (Chapter 6). Didactic remarks as well as different types and numbers of *chair interventions* are presented and explored in Chapter 7. A limited follow-up study is performed over a two-year period with a small group of participants to search for changes over time is described in Chapter 8. Finally, Chapter 9 offers reflections from several perspectives on the research results achieved.

2 Characterisations of Philosopher's Activity*

1 Philosophy, philosophising and being wise

This thesis aims to assess the philosophical quality of thinking patterns that are exploited by individuals and, in particular youngsters, during group discussions. 'Philosophical quality' strongly relates to other concepts, i.e., 'philosophy', 'philosophising', and 'wisdom'. A common understanding of these terms is thus needed to describe philosophical quality or a potential philosophical talent. In this chapter, non-specific general features of philosophy and of the activitity of philosophising are searched for.

Historical perspective

'Philosophy' has evolved through history. Philosophical concepts and philosophising emerged in the city of Athens during the fifth century B.C. because of favourable political influences and intellectual environment (Hadot, 2003). For example, the meeting of the Athenian legislator Solon with Croesus, king of Lydia (ca. 700 years B.C.) is described by Herodotus (1996) as a philosophical activity. Herodotus considers Solon's love for wisdom (philosopheon) as a drive to explore and observe the world. According to Hadot (2003), the noun 'philosophy' and the verb 'philosophising' refer to an exploratory life style, and to existential choices directed to the competence of 'areté' as a prospect of virtue, excellence and wisdom. By Aristotle's time (384 - 322 B.C.), 'Philosophia' also referred to science in general and could be pluralized depicting several branches of science.

Philosophy and wisdom

This perspective suggests reciprocity between the concepts 'philosophy' and 'wisdom'. For ancient Greeks, philosophy and philosophising involved collecting knowledge and experience. This could be achieved by travelling, collection of encyclopaedic knowledge, or by emphasising a competence and attitude in life. For the Presocrats, the lust for knowledge and experience evolved into a drive or commission to investigate and rationally colour the world. In this tradition, philosophy was incorporated into a life style dedicated to reaching the competence and virtue of 'areté' and wisdom. It was the task of the society to educate gifted youngsters in philosophising. As the victory of democracy was proclaimed, all citizens were qualified to devote their lives to philosophy, to the love of 'beauty' as virtuousness, and to attainment of wisdom. Paradoxically, wisdom was also considered as an ideal that could never be reached. Philosophy was the preparatory exercise to gain wisdom. The ancient philosophers did not conceive themselves wise: they did not know what was and was not, and were never sure about what had to be done or avoided. They pursued wisdom asymptotically without ever

* The main part of this chapter was published in Teaching Philosophy, March 2000.

reaching their goal. Wisdom posed as an ideal that attracted and guided philosophers (Hadot, 2002). If their exercises towards wisdom were successful, the process of philosophising was finished.

Philosophy as life style and the Socratic dialogue

Possibilities for philosophising were universal and ubiquitous in daily life experiences. Doing philosophy always took place in school environment with exercises that were conducted through physical, discursive, and contemplative activities. The discursive exercise in language was the most significant method. Philosophy, as an exercise in arguing, was part of a life style in classical education. This life style was designed to develop several skills and competences in the arts, sciences, administration of justice and law, and politics.

Socrates' reasoning and arguing was integral part of his life and of the way he died. His attitude and actions carried as much philosophical meaning as his words as exemplified by 'Socratic dialogues' – exercises in thinking and life style based on questions and answers disinterestedly pursuing the truth through analytical discussion. Participants would discuss pre-determined answers to philosophical questions and discovered knowledge independently by themselves. In this play of asking-and-answering, philosophers first secured the mind against sole solipsism, and then refrained from vested and individual interests. Philosophy transcended itself perpetually into a 'logos' common to all participants. Presuppositions of conventional knowledge were unmasked by questioning and brought up for discussion. It was in the Socratic dialogue that philosophy and the philosophising fraternity arrived at their typical meaning. Plato described Socrates as throwing his participants into confusion (aporia) and raising doubts by his naïve attitude and seemingly innocent questions. It is here that philosophising demonstrates its ironic and tragic elements. Irony refers to the philosopher knowing that he does not know. Philosophy inclines towards indecisiveness and uncertainty. Irony also concerns the position of being simultaneously inside and beyond reality. Tragedy refers to the philosopher tortured by the desire for wisdom that slips away time and again.

Philosophy through time

In ancient Greece, philosophy as a life style was characterized by perpetually coping with philosophical questions and their application to a philosophical way of life. Through time, the role of philosophy has changed by the rise of sciences and by the tendency of philosophy to be guided by sciences. Comprehensiveness, engagement, and disinterestedness faded and had almost disappeared by the end of the 19th century. Modern philosophy focuses on the exploitation of specialised thinking patterns and is seen more as a theoretical activity that gives birth to an academic domain with its own rationale, detached from other disciplines. It examines the construction of knowledge and generates thinking patterns and notions referring to typical philosophical questions (i.e., questions investigated systematically by philosophers). Life styles concerned with engagement are now the domain of

religion, whereas encyclopaedic knowledge is the concern of a variety of specialised scientific disciplines. In this respect, philosophy is the mother of all sciences.

Beside the academic domain, and outside the scope of this study, a more popular concept of philosophy emerged during the 20th century, depicting general and particular trains of thought like fantasising or management styles. The verb 'philosophising' is commonly used with reference to less rigorous forms of thinking. Concepts of philosophising and wisdom diverged with philosophising viewed as exploiting thinking patterns to conduct philosophy, while wisdom became a cognitive stance or attitude and an object of psychological investigation. Today, philosophy as 'thinking patterns' has become detached from philosophy as 'posture' and as 'life style', and from wisdom as competence and as attitude; and finally from other specialised scientific disciplines.

Modern philosophising youngsters

Nowadays, doing philosophy is no longer exercised like in classical times, without direct interest. Classical philosophising may revive only in discourses when philosophical questions are disconnected from vested interests or selfish context: for example when young people not yet locked up within the limits of a knowledge system, are engaged in philosophical questions as a form of mental play. Childrens' radical way of reflecting and jumping to conclusions, their unselfishness, and the fluent transitions between their thinking and acting are reminiscent of classical philosophising exercises. Socratic irony may even be re-valued in such exercises (i.e., knowing not to know, humour, being simultaneously inside and beyond reality). The way youngsters cope with philosophical questions seems analogous to the classic philosophical life style. The same may hold for the elderly, detached as they are from business, social relations and the material world.

In recent years, attempts to define or characterise philosophy have varied widely. At the same time, considerable unanimity exists among philosophers concerning the boundaries of the 'discipline'. Apart from extreme positions, philosophers will normally agree about questions of what comprises a philosophical work and what does not. Concentrating on philosophy as an activity, this chapter will construct a synthetic view from characterisations given by investigators with various philosophical orientations. Common philosophical denominators may be found in joint historical threads and in similarities between modern philosophical thinking patterns.

Aiming to find non-specific and general characterisations of the acitivity of philosophising, an inventory of ideas, and biographic aspects of key philosophers, will be illuminating. Many contemporary philosophers have passed individually under review and present a more or less consistent view in their explicit description of characterisations of philosophy. In this study, the search for common features of philosophy is guided by 1) well-known 20th century philosophers, 2) with explicit views on philosophising, 3) originating from diverse orientations. The search will be continued until the rising selection of characterisations is saturated and more descriptions will slip back into repetition of earlier derived characterisations.

2 Characterisations of philosopher's activity in the 20th century

Fortunately, explicit descriptions of philosophising were passed on the 19th century covering a broad spectrum of views in western philosophy, ranging from the philosophy of science to those contemplating existential experiences. This study will restrict itself examining the development of philosophy over the last 100 years for saveral reasons. First, 19th and 20th century philosophy incorporated and built upon developments in the past while developments in sciences have changed the concept of philosophy thoroughly. Furthermore, the emphasis in this study is on philosophical thinking patterns of modern youngsters. Like contemporary philosophers, they think in a current paradigm and use language that refers, more or less, to similar objects and situations. As will be demonstrated in Chapter 3, some children show thinking patterns that seem to fit in earlier periods of philosophising. Common to all characterisations of philosophy, is its conception as an autonomous cognitive activity that abstains from external means like instruments, specialised knowledge or appeal to authority. In all traditions, philosophy is regarded as the independent production of consistent thought avoiding dogmas, certainties or definitive judgements. This approach is followed in the present study, which perceives philosophy as the mental activity of individuals and groups engaged in philosophical questions. It includes an attitudinal stance of being, and at the same time, maintaining an intellectual distance to philosophical questions. As characterised briefly by the following literature selection a wide range of beliefs and mental activities seem to be associated with philosophy. Common in these is the conception of philosophising as a rational activity, not based on any empirical investigation. Adding nothing to our knowledge, the activity of philosophising concentrates on the analysis of concepts, explanation of notions and beliefs, elucidation of presuppositions, and clarification of everyday experiences.

According to Deleuze and Guattari (1994), philosophy is the creation of concepts rather than contemplation, reflection or communication. Philosophy is the opening that allows thoughts to escape from the constraints that seek to define and enclose creativity. These philosophers stress philosophy in its capacity of knowledge through pure concepts. The concept that is formed, invented and fabricated is the contour, the configuration, and the constellation of an upcoming event. When creating concepts, the task of philosophy is always to extract an event from things and beings, to set up the new event, or the possible as events.

Nozick (1981) stresses at least two characteristics of philosophical thinking: reasoning and search for alternatives. Although the quality of a philosophical theory is measured by the decisiveness of its arguments, the value of arguing and proving by itself is subordinate to philosophy in general. Philosophical arguments are means to clarify a view; which is not identical to arriving at a consensus. Philosophical discourse is the search for explanations. Philosophers therefore have a permanent interest in paradoxes and uncertainties, and philosophical theories always remain tentative hypotheses. Philosophy is positioning situations in a network of alternative possibilities.

The investigation of presuppositions concerning our view of reality involves the exploration of the boundaries and structure of our experience. Husserl (1968, 1996) recommended suspended judgement or 'epoche' with regard to the existence of objects of consciousness to explore the nature of experience. We must suspend or 'bracket' our natural attitude towards the world in order to describe the essential structures of experience. This involves suspending our scientific presuppositions and our practical engagement in it.

Following on Husserl, Merleau-Ponty (1967), attempts to elucidate the structure of our experience. In his view, all knowledge is infected by an ambiguous mode of existence given in our perception, in our body, and in our language. Because of this fundamental ambiguity in our relation to the world we will never acquiesce in definitive conclusions. And because we will never obtain a complete knowledge of the world, thinking and philosophical discourse will never stop. Primarily we perceive. Instead of knowing, we believe, and instead of analytic thinking by splitting up experience into smaller constituents, (e.g., sensations and qualities) Merleau-Ponty promotes the idea of a radical reflection. Our thinking expresses itself in words, and the value of words is always temporary because words will never adequately represent an ever-receding reality.

According to Sellars (1963), philosophy is 'the eye on the whole'. The aim of philosophy is to understand how 'things' hang together in the broadest possible sense. Sellars refers to Husserl's 'bracketing' to elucidate his conception of 'reflective knowing'. To be able to think is to be able to measure one's thoughts by standards of correctness, of relevance, of evidence without acquiring new knowledge. It is impossible to learn conceptual thinking by being told the rules; and whatever else conceptual thinking makes possible, it does so by virtue of containing a way of representing the world. To achieve success in philosophy you should 'know your way around' in the order of things. The emphasis is on conceptual clarification.

Critical thinking also implies identifying differences and contrasts. This conforms to the ideas of Derrida (1967). His key term *differance* refers to the process of giving meaning to a word; the discovery of new denotations and conditions of making sense in the spirit of De Saussure (1975); and also to the postponement of definitive judgement (deconstruction of meaning). Like Merleau-Ponty, Derrida analyses the meaning of experiences. His position is highly sceptical. To know an entity is to situate it within a frame of relationships and oppositions. Deconstruction is a way to overcome the consequences of totalitarian thinking and of the formation of definite judgements.

The linguistic turn, characteristic of modern philosophy, did not really change the objectives of philosophical research. Classical problems still form the central focus of philosophical thinking. But from now on, these will be interpreted in the relationship between language and reality. According to Wittgenstein (1969, 1976), the sole function of philosophy is to guard the boundaries of meaningful language, to elucidate philosophically problematic sentences, and to show that attempts to transgress the boundaries of meaningful language are futile. Since sentences are materialised thinking patterns, the philosophical exploration of concepts is equivalent

to that of linguistic utterances. According to Wittgenstein, philosophy is not a science (*Wissenschaft*). There are no philosophical propositions, nor is there philosophical knowledge. The task of philosophy is conceptual clarification; its goal is not knowing, but understanding.

Finally, we focus on the ideas of two philosophers who were very influential in the field of education: Nelson (1970) and Dewey (1956, 1966). For Nelson, philosophy – although it produces no new knowledge – is concerned with truth. The philosophical truths belong to the structure of reason itself and so can be discovered by investigating the presuppositions of one's own experience, in principle by 'looking inward'. We can never be certain that we did not deceive ourselves. And we can never be certain that we went as far as we could go. Here, the Socratic method enters: If we can reach consensus in a discussion in which all participants in cooperation try to establish the truth of the matter, it is reasonable to assume that we are on the right path even if insights gained remains open for later revision. Complete clarity regarding the conceptual relationships involved and consensus are principal conditions in order to reach truth.

Although Dewey's philosophical views stem from quite another philosophical orientation, there are striking similarities between his view on philosophical education and Nelson's. Both stress interaction or discussion as intrument to philosophise. Dewey's ideas about the activity of philosophy were based on a theory of inquiry, a general account of how thoughts function as merely useful instruments for managing, ordering, and anticipating the observable world. In his view, no knowledge-claim, no moral rule, no principle, or ideal is forever certain (epistemological and moral fallibilism). There are no ready-made answers, there is no way to apply formulae to derive a correct solution and no way to definitely prove that a proposed solution is correct. The process of problem solving is governed only by the need for a solution to the problem. Development of, for instance, ideas about morality is welcomed. Interaction is used to construct a contemporary reality, whereas it serves truth finding according to Nelson. Judgements based on reflective thinking, sensitive for multiformity and new ideas are more likely to be valid and insightful than beliefs derived from authority, emotional commitment, or narrow reasoning. Dewey's views on philosophy and education have been a major contribution to the background theory of philosophy for and with children (Lipman, Sharp & Oscanyan, 1977).

3 Main features of philosopher's activity

The process of philosopher's activity must be distinguished from its products. All presented insider views concerning the process of philosophising stress the absense of acquiring new knowledge, the analysis of concepts, explanation of beliefs, elucidation of presuppositions, and clarification of everyday experiences. Considering these characteristics more detailed, they can tentatively be categorised into three aspects of philosophical thinking, although emphasised differently in the presented views: one that stresses analytical and reasoning qualities, one dealing with ambiguities, vaguenesses or borderline explorations, and one stressing the contact with real life experience.

The first aspect concerns analytical and reasoning qualities. These imply skills like consistent thinking, correct reasoning, problem solving, categorising, clarification and understanding of identities in the structure of relationships and oppositions. This may even be true for some types of moral reasoning and reflective judgement (Kohlberg 1971, Kitchener 1990). These skills can be labelled as informal logic, concern mainly convergent thinking patterns and have been extensively studied.

The second characteristic concerns dealing with ambiguities and vaguenesses. This tendency does not refer to pure cognitive matters. Precise standards of correctness do not exist for qualities of this second kind. Dealing with vagueness and ambiguity does not have specific outcomes. Moreover, these tendencies refer to an ancient stance of philosophy as something that is indecisive and uncertain, to an ideal that is never attainable, and products that are never final. These qualities create problems when measuring performance in an area that, by definition, deals with uncertainties and un-ending conceptualisation. Creativity research in the 1970s exhibited similar measurement problems (Brugman & Dudink, 1976, 2002). Searching for and dealing with ambiguities, vagueness, paradoxes and uncertainties, the suspending of judgement, and the monitoring of the boundaries of meaningful expression, are all tentative mental moves essential to the philosophical quality of thinking patterns. These qualities also contain cognitive and attitudinal elements. Recognition of vagueness represents the cognitive element, while accepting them denotes the attitudinal element that also makes contact with the ancient tendencies. Both are analogous processes and require capacities like being honest, free of charge, and having a fundamental openness to the unknown. These capacities and attitudinal elements of acceptance and interaction with the experience are mirrored in events that are studied in wisdom research. This second tendency is sometimes best characterised by emphasising what philosophy is not: it does not deal with certainties, definite judgements, ready-made answers, a single definite way of thinking and knowing; it does not call upon authorities or emotional commitment, and produces no uniformity of belief. Russell in his *History of Western Philosophy* characterised the task of philosophy as follows: "To teach how to live without certainty, and yet without being paralysed by hesitation, is perhaps the chief thing that philosophy in our age can still do for those who study it". (Russell, 1974; p. 14).

The third aspect of philosophical thinking refers to concrete experience and is connected to the meaning of real life. Although thinking patterns mainly refer to a ratio that is supposed to function outside of the experience, the discipline stresses the smooth and flexible transfer from abstract thinking levels into concrete occurrences, and from these into several mental representations. Once this transition is interpreted as an act of translation, it is then referred to as an act of abstraction, or of giving concrete forms. In this respect, comparisons must be emphasised. To compare two events is to search a common level covering both instances in order to detect similarities and differences. This requires being able to automatically switch between different levels. The philosopher has to transpose a set of relationships from the one domain into another. Philosophical thinking patterns also reflect a capacity to move such sets of relationships from the experience to abstract thinking patterns. Qualities like Merleau-Ponty's 'radical reflection' or Nelson's 'regressive abstraction' point to this tendency. Philosophy

is more than analysing or speculating, it concerns the conduct of life itself and so presupposes a readiness to reflect on, and clarify experience. This implies not only cognitive skills, but attitudes as well.

As a consequence, evaluation of performance in philosophy cannot restrict itself to specific, describable cognitive skills or to the identification of specific outcomes. Identifiable manifestation of doubts, uncertainties, ambiguities, and reflections on life experiences must be detected. To consider these features as an object of study is to lead the restricted philosophical matter into a wider conception where the Humanities become visible. Attitudes and capacities reflecting specific human kinds of reaction on experiences and of interaction with the environment, skills or competences to deal with uncertainties are also subjects of psychological research. Where wisdom in philosophy is an asymptotically deferring ideal that guides and attracts, in psychology it is the tangible object of empirical investigation, found in the real world and defined by concrete qualities. To detect indicators for philosophical qualities in broad sense, the field of empirical wisdom research by psychologists must also be explored.

4 Wisdom: psychological approaches in the 20th century

Wisdom is a form of understanding that unites reflective attitude and practical concern (Kekes, 1983). While understanding and a reflective attitude are philosophical qualities, practical concern is not seen as part of a purely cognitive activity or of the academic discipline of philosophy. Many descriptions exist of concepts of wisdom as a research topic, although over time clear shifts in these concepts may be observed. In the western tradition, wisdom evolved from knowledge about rules of proper conduct, areté and a weak reflection of divine wisdom into the cynical cognitive stance found with Nietzsche (Brugman, 2000).

Today, wisdom is characterised by traits like overall competence, good judgement and communication skills, being able to see things in large frameworks, and exceptional understanding (Chandler & Holliday, 1990). In lay opinions, wisdom may be termed 'sagacity' and refers to an attitude towards knowledge. Most systematic psychological research on wisdom in the 1990s focussed on connections between cognitive or reflective thinking patterns and practical concerns. Some studies, like those of Baltes & Smith, adopted a pragmatic stance, while others focussed on an epistemological approach that emphasised conceptual and empirical investigations (e.g., researchers like Kitchener, Sternberg, Riegel and Meacham). Possible relationships between wisdom and intelligence, wisdom and education, wisdom and personality traits, and wisdom and age were also examined.

Pragmatic stance

Baltes & Smith (1990 b) define wisdom as a form of crystallised intelligence and as good judgement in practical matters of life (like Aristotle's 'phronèsis'), particularly complex matters that lack proper problem definition and solutions.

Their Berlin-school developed procedures for measuring features of wisdom and formulated five cognitive criteria based on expertise in the pragmatics of life: 1) wide factual knowledge; 2) rich procedural knowledge; 3) life span contextualism; 4) relativism; and 5) uncertainty. This conception mainly stresses the connection between thinking processes and life experiences. More recently, Baltes & Staudinger (2000) expanded their concept with a clear moral component and a spiritual point of view.

Epistemological stance

In defining wisdom, researchers like Riegel (1973), Kitchener (1990), Sternberg (1990, 2000, 2003) and Meacham (1990) focus on cognitive epistemological stances. Kitchener (1990) characterises wisdom as an intellectual awareness of the limitations of knowing as a result of solving poorly-defined problems. She describes the development of epistemic cognition into a reflective judgement in seven stages. Epistemic cognition implies the possibility of monitoring the solvability of a problem under any condition. Wisdom is consequently characterised as the awareness of a never-ending process to arrive at penultimate answers. During this process, the researcher must always select the best possible judgement. A judgement is wise if it reflects recognition of the limits of personal knowledge, acknowledgement of general uncertainty, and humility about one's own judgements. This conception coincides with the second of the main philosophical features: the view that definite standards of correctness and outcomes do not exist.

Tolerance for and construction of new uncertainties, in addition to doubts and questions about what might be known, complies with the views of Meacham (1990) and Brugman (2000). According to Meacham, wisdom is the middle course between knowing and doubting. It is a constant factor. The essential element in wisdom does not concern what is known, in pure knowledge, in beliefs, or in values, but in how knowledge is held and put to use. Wisdom is an attitude and an awareness of the fallibility of knowing, originating in interpersonal relations. Brugman perceives wisdom as expertise in uncertainty, encompassing metacognitive, affective and behavioural components. Generally, metacognition deals with deliberate guiding of one's own thinking. Here, the metacognitive component deals with uncertainty on the possibility of gaining solid knowledge about reality, jeopardising fixed beliefs. The affective component denotes or relates to a missing emotional disturbance in view of uncertainties. The behavioural component refers to the fact that these uncertainties do not lead into inertia. The wise man still can make clear decisions to act.

Intelligence and wisdom

When it comes to developing epistemic stances, intelligence is the most salient and best measurable cognitive capacity. Cognitive qualities attributed to a wise person refer mainly to intelligence (e.g., intellectual awareness, judging, recognition, knowing and doubting, dealing with uncertainties). How are wisdom and

intelligence, the cognitive quality par excellence, related? Sternberg (1990, 2003) points to their difference. The metacognitive stance that he formulates refers to the relationship between thinking and acting in one's attitude towards knowledge. A wise person tries to understand situations and events vis-à-vis their presuppositions, meaning and limitations. This reminds one immediately of Sellars (1963) and Derrida (1967). The wise person will use his wisdom in various ways depending on what is valued within a given environment. This deviates from what might be seen as intelligence. Where the intelligent person tries to eliminate ambiguities and tries to break down barriers or overcome obstacles within conventional frameworks, the wise person welcomes ambiguities, tries to understand these, and enjoys investigating such barriers or obstacles. This way of dealing with cognitive skills, with attitudinal stances, and playing with the interconnection of thinking and real life experience is in full harmony with the three pillars of philosophy. Sternberg (2003) developed an interesting dialectical conception on wisdom: wisdom is the synthesis, following intelligence as thesis and creativity as antithesis.

The epistemological approach to wisdom cannot entirely be characterised by standard models of cognition and intelligence because they mainly concentrate on analysing, reasoning and problem solving. Consideration of philosophical thinking patterns shows the same negotiable restrictions. The connection between thinking patterns and cognitive development will be discussed in more detail in Chapter 3. Relevant qualities also focus on processes of problem selection through interacting with the external world. Riegel (1973) emphasises the continuing dialectical character of these processes. Despite their development in stages, such processes will never terminate but rather culminate in a dialectical train. At the end of each stage, one is searching the controversies transcending that stage. This continues after all stages are over so that thinking patterns remain open to ambiguity, contradiction and uncertainty. A wise attitude toward knowledge therefore implies the casting of doubts on issues that seem to go without saying.

The dialectical character of thinking patterns, as stressed by Riegel and Sternberg, particularly reflects interaction and dynamic exchange between abstract levels of thought and concrete experience, in accordance with the third main pillar of philosophy. This interaction is tolerant of unknown, non-fixed and non-fitting experiences. This also holds for incongruences between instances of concrete and abstract cognitive levels. Antoine de Saint-Exupéry (1980) exemplifies this point in Chapter XV of *The Little Prince*: a scientist states to know everything in his field, discipline of geography, but is unable to apply such knowledge in practical experience or to link these experiences to his encyclopaedic knowledge in his books. This is metaphorically demonstrated by the way Saint-Exupéry's expert records 'certain' knowledge in ink and 'uncertain' knowledge by erasable graphite. Furthermore, the author links the uncertainties of experiences to the vulnerability of scientist's mental state. In drunken condition, the geographical explorer observes everything in double and, as a consequence, records things twice. As uncertainties refer to non-definite and temporary knowledge, Saint-Exupéry's geographer refrains from noting blooming flowers in his book of 'certain' knowledge. As he is unable to link ever-lasting knowledge and temporary experience, the geographer

cannot expand his knowledge and further explore his surroundings. As a consequence of his inability to interact with or to connect theory with real life, the expert cannot develop into a wise man and is doomed to stay a scientist. This story teaches to value openness to uncertainty over closed encyclopaedic knowledge, where wisdom is concerned.

Education

Wisdom and intelligence are correlated, but not above a certain threshold value of intelligence (Brugman & Dudink, 2002). Psychometric intelligence and wisdom are positively correlated with educational level (Brugman, 2000). Acquired knowledge is crystallised intelligence gained through education. All intelligence, fluid and crystallised, trained or exploited, can be connected with levels of education and with age. Crystallised intelligence is cumulative and will therefore increase through lifetime, given intact brain functions.

Several researchers describe capacities of wisdom as distinct sets of processing operations (Guilford, 1988). Such operations permit individuals to solve problems, to create products, and to discover new knowledge in a divers array of culturally valued activities (Berk, 1997). Gardner (1999) proposes seven independent intelligences applied in culturally meaningful activities: 1) linguistic; 2) logico-mathematical; 3) musical; 4) spatial; 5) bodily/kinaesthetic; 6) interpersonal; 7) intrapersonal intelligence; and provisionally, 8) existential intelligence. In particular, the last two intelligences (intrapersonal and existential intelligence) seem relevant with respect to wisdom. Intrapersonal intelligence comprehends processing operation to discriminate complex inner feelings and to use them to guide one's personal behaviour; knowledge of one's personal strengths, weaknesses, desires, intelligences and qualities of wisdom alike. Existential intelligence (in myth, art, science and philosophy) refers to the a separate intellectual domain focussing on thinking about the meaning of life and human existence. Although conceptually sound, not much empirical support has been found for this kind of intelligence to date. It is possible to further extend the specific domain of intelligence into theories of education. The educationalist, Hirst (1980) describes philosophy as a separate knowledge domain. He distinguishes seven knowledge domains: mathematics, sciences, humanities, history, religion, literature, arts, and philosophy. Each knowledge domain encompasses its own concepts, its own rational beliefs, its own ways of reasoning, its own ways of understanding events in perspective of its presuppositions, meaning and limitations, and its own forms of justification (Ritzen, 2004).

Creativity

The significance of creativity in wisdom is emphasised by Sternberg's wisdom as synthesis, following intelligence as thesis and creativity as antithesis. Creativity connects openness to the unknown and production of the new. However, so far no empirical relation between these two concepts has been confirmed. Many

creativity tests are reported to have problematic validity (Sternberg & Lubart, 1996). In his attempts to study creativity through psychometric tests, Guilford (1988) found divergent cognitive activities like creativity, positively correlated with convergent intelligence. However, no such correlation appears to exist when IQ exceeds 120. Many creativity tests focus on associative thinking patterns, which is in accordance with the third tendency of philosophical thinking. Simonton (1994) found that creativity blooms during periods of revolution or oppression and that creative persons often originate from unhappy families, who have suffered from early parental losses and traumatic youth. In his 2003 paper, Simonton shows that creativity is more manifest in products of scientists, philosophers and artists who incorporate wide spectra of scientific domains than mono-disciplinarists. Interdisciplinarity evokes tolerance for ambiguity, openness to the unknown, searching the dialogue and opposition. The creative person is independent, exceeds the conventional, and questions authority and all that is self-evident (Simonton, 1994).

Personality traits

According to Staudinger et al. (1997) and Brugman (2000), wisdom can be positively correlated with two of the 'Big-Five' personality traits. These traits were originally formulated as 'openness to experiences' and 'emotional stability'. Despite its general recognition, openness to experience is often identified as a problematic factor. In a philosophical, inquisitive and insightful sense, Saucier & Goldberg (1996) found some connection between this factor and autonomy. De Raad (1994) identifies openness to experience (O) as 'intellectual autonomy'. In her 'Construction of the Five-Factor Personality Inventory', Hendriks (1997) found lexical representations of this factor (O) with items like 'links facts together', 'wants to form his/her own opinion', and 'analyses problems'. The first item refers to reasoning skills and the connection between thinking and experiencing; the second to autonomous production of thoughts; and the third to reasoning qualities. The negative pole of this factor was loaded with items like 'follows the crowd', 'copies others', and 'does what others do'.

Wisdom and age

Opposite views exist on the relation between thinking and growth with age. As wisdom is linked with innate personality traits, no dramatic changes in wisdom development may be expected during life, in contrast to public opinion that wisdom grows with age. If wisdom is seen as a kind of cognitive development that is connected with crystallised intelligence, it might grow through a life-time. However, no such relation has been shown to exist in studies on fluid intelligence, creativity, and expertise in uncertainty. Many empirical studies with a diversity of wisdom operationalisations have been undertaken, but a relation between wisdom and aging has yet to be found (Baltes & Smith, 1990b; Birren, 1990; Ardelt, 2000; Brugman, 2000; Bruman & Dudink, 2002). These studies refer to developments in adults 30 years of age and older. However, as a form of dialectical maturity,

wisdom may be accessible at all ages. Meacham demonstrated positive empirical evidence for wisdom potential in childhood, adolescence and young adulthood (Sternberg, 1990). For example, he shows how at age 15, Anne Frank (in her diary, *Het Achterhuis*), moved from a position of confident knowledge to a moderate and wise position between knowing and doubting. In contrast, aging may induce loss of wisdom as the current intellectual climate prevents adults from spontaneously expressing their prime ideas, forces them to defend their views too early and to abandon tentative notions rather than being permitted to playfully entertain ambiguous or contradictory positions. Other potential threats to wisdom are confrontations with rapid technological changes, cultural changes, and personal tragedies. So, if the character of philosophical dialogues with youngsters belongs to wise and philosophical thinking patterns, such threats will inevitably leave their mark.

5 Main features of philosopher's activity and wisdom as thinking patterns

Correspondences

Wisdom and philosophising as expressed as thinking patterns and embodied in mental acts and mental states, can be observed and then, approached empirically. The main features of philosophy are strongly fed by those of wisdom, stressing reasoning and analytical qualities, and abilities to frame events in wider abstract or concrete contexts. They include tolerance for ambiguity and uncertainties, and a penchant for dynamic interaction between knowing and experiencing. Expressions of 'not being sure' represent a sense of ambiguity, as do the opposites, 'it might be otherwise' or 'it might mean different things'. Autonomy and wonderment as the construction of the unknown are examined in both philosophical and wisdom research. The involvement in life pragmatics is also considered essential. All these features contribute to the foundations for recognising philosophical quality expressed in thinking patterns. Although philosophy pays less attention to attitudes and affective matters than is done in wisdom research, a substantial agreement in basic requirements is promising if clearly phrased indicators for detecting philosophical quality are to be found.

Discrepancies

Investigations into wisdom and into philosophical thinking patterns differ fundametally in their approach. The philosophical approach focuses on overviews stressing the formation of trains of thought, whereas the psychological approach deals with individual expressions. By studying philosophy and its performers, general forms of attention attracting problems are highlighted and associated thinking patterns can be recognised. Philosophically qualified thinking patterns are usually detected through their topics and developments in these patterns are judged thematically. Philosophical estimations are from the overview level, and independent of time and individual contributions. The psychological approach,

on the other hand, focuses on individual contributions to tackle problems, in the course of which, thematic ends are presupposed. Wisdom places human beings and their way of coping with experiences in a central position, while philosophy deals with conceptual problems and ignores the human being. Limitations of the philosophical approach should be overcome by psychology, and the shortcomings of the latter might be counterbalanced by a philosophical overview. Unfortunately the two approaches seem not compatible and a common denominator is needed to bring incomparable concepts of epistemology and cognition, of empirical findings and conceptual conditions, skills and aptitudes together. Closing the gap between philosophy and psychology will be discussed in Chapter 3.

Philosophical quality materialised in indicators

Indicators must be developed to judge thinking patterns in the absence of definitive results or definite answers. Manifestation of doubts, uncertainties, ambiguities and reflections on life experiences belong to philosophical qualified thinking patterns and can be detected in verbal and non-verbal expressions. Analysis and reasoning can also be observed in verbal expressions. Smoothly switching from abstract to concrete and vice versa can be found in real life discussions between youngsters too. Recordings of such discussions can serve as a data base for judging the presence of specific elements indicative of the main features of philosopher's activity. However, recognized philosophical qualities should be based on more than just static entities in verbally expressed thoughts. Philosophical quality is a capacity that is realised in individual expressions, but also on a level that transcends these, reflecting a developing train of thought, an awakening to questions, uncertainties and new findings. Elements of philosophical quality need to be detectable at multiple levels: in single expressions and in trains of thought, both on an individual level and in groups of philosophising youngsters. Analysis of philosophical quality should use indicators and combinations of indicators derived from the main features of philosophy and wisdom, but also by estimating the exchange of thoughts as a whole.

In the joint-venture dialogues of this study, as in Socratic dialogues, the philosophical quality is hidden in undivided inquisitive thinking operations, built by more than one individual. The philosophical quality of such a dialogue as a whole is based on the contributions of dominant individuals and dialogical events that lead to a thematic philosophical outcome. A surveying overview is required to bring out the main philosophic features in the context-rich whole because relationships within group discussions need to be judged. The production of autonomous lines of thought is detectable by placing them in the larger framework of common sense thinking. To determine the intellectual distance to the philosophical question, judges must compare lines of thought with the starting-point or initial question. Increasing numbers of questions and uncertainties, and a growing consciousness of the degree of complexity, can only be observed at an overview level. The same holds to the quality of persistence in systematic exploration. Without the condition of receptivity for systematic inquiry, philosophy would degenerate into interminable thinking patterns. Such philosophical qualities should be identified in detectable

elements of joint ventured dialogues. This bi-level judging requires development of philosophical quality indicators that are correspondingly valid for both levels, and will be discussed in Chapter 4.

Expectations regarding correlates of philosophical indicators

Differences between young participants must be documented in a consistent, reliable way. If knowing and experiencing are significant components in philosophical dialogues, we may expect some correlation between these qualities and philosophical quality. Sternberg's (1990) studies on wisdom show complex relationships between cognition and thinking patterns reflecting wisdom qualities. On the one hand they contain reasoning skills, on the other they refer to a separate domain. Some convergence between a measure of philosophical quality and a measure of general inelligence is expected, as well as a considerable divergence, demonstrating their distinctiveness. Level of education and philosophical quality may be correlated positively, as level of education is linked with knowledge, crystallised intelligence, and processing operations. Philosophical quality should be connected with intellectual autonomy as a personality trait on conceptual grounds. Whether philosophical thinking patterns and age are correlated depends on what is emphasised most: crystallised cognition or creativity. While cognition grows with age, creativity may be infected negatively by conventions and experiencing. As experiences matter, some correlation may be expected between philosophical quality and regularity of life course. All these theoretical expectations form a substantial part of the so-called nomological network around the construct 'Philosophical quality' (Figures 6.1 and 6.2).

6 Summary and discussion

Three main features of philosopher's activity were uncoverd as common characteristics:

1. Analytical and reasoning qualities
2. Qualities detecting ambiguities, vagueness, uncertainty, and borderline explorations
3. Qualities realising a wide framework with connections between theory and practice or real life experiences.

The present project aims to measure philosophical quality. Empirical psychology has much experience with the measurement of thinking patterns. Empirical psychological studies into wisdom focus on similar qualities as in the present project. The measurement of an elusive construct as 'wisdom' gives confidence that a psychometric approach to philosophical quality may be succesful. The data source for such approach is found in recordings of discussions on philosophical topics. The verbal and non-verbal expressions of discussing youngsters are considered as materialisations of thinking patterns. Theoretical considerations presented allowed the formulation of expectations concerning relationships between a measure of philosophical quality and measures of intelligence, educational level, autonomous thinking, and age.

One paradox remains. Philosophising with youngsters in an educational context seems rather contradictory. The school context presupposes clear educational or instructional aims, while at the same time philosopher's activity asks the individual to maintain openness to the unknown and to never reach a final answer.

How can we expect to find an index for philosophical quality? Specific thinking patterns, like the autonomous production of consistent thoughts on reality, may be trainable to some extent and influenced by controllable circumstances. But is there a well-founded reason for it? Maybe it is our intuition that tentative interpretations of life events may encourage intellectual curiosity and may prevent children to accept things that 'go without saying'. Thinking in terms of 'what else is possible' and being open to views of the unusual, the unfashionable, the uncommon, or even 'sick', may have social significance, serving an emancipator purpose and enlarging the democratic content of society. While standard education and decorum presuppose some self-evident structures that undoubtedly serve an instrumental good, they also paralyse fundamental discussions about the innumerable possible interpretations of reality. Even in philosophy, the tendency can often be to accept the safe and prevailing position. But the characterisation of philosophy is based on openness, the search for oppositions and borderline cases, vagueness and ambiguity; it does not seem to be governed by rationality alone. Perhaps it is just this sensitivity for vagueness and ambiguity that is a vehicle for development and fundamental creativity, because it is able to strip concepts of their rigid definite meanings. It is the challenge of this study to find indicators that satisfy the paradox: no development without vulnerability.

3 *Thinking Patterns* – Bridging philosophy and psychology

How can philosophical qualities that are elusive by nature be mapped and measured? Concepts that are difficult to measure, for example, intelligence, are examined in psychology through the use of empirical methods. Is an empirical approach to philosophy possible without influencing its intangible nature?
Let us first look at children's philosophising. Children ask questions, they reason, formulate arguments, and construct new suggestions. They discuss themes that are philosophical by nature in a philosophically interesting way. Philosophers may identify such thinking patterns as philosophical, taking them at face value. These thinking patterns give an insight into their quality, regardless of their cultural heritage, standards of correctness or level of maturity. Sometimes these thinking patterns can be justified in terms of agreed-upon standards, sometimes not.

Rosa (ten years old) was asked: 'What actually is a horse?' Rosa: 'A horse, well, that is all you see after the birth of a horse.' On my question what that exactly meant, a much more conventional answer followed: 'A horse is an animal with four legs.' An animal with three legs, and even one without legs, appeared to be a horse in her eyes too, because a horse finally is the upper part. 'Which part of the animal should basically be left for determining it as a horse?' This question seemed simple: 'The head,' she answered shortly. 'But if you would cut the head from the body, what exactly is the horse, the head or the body?' Now Rosa dropped silent, obviously pondering the question. Then, to my astonishment, she concluded that it is the shit that would matter: 'nothing else is shitting like a horse. If his head is off you note by the crap that it has been a horse. So it is a horse.'

In this example, Rosa gave evidence of her genuine own 'logic'. Entrance and exit of the animal are connected. The horse is the result of what is flowing through the animal. Is this a philosophical thinking pattern? In any case, it is a succession of thoughts that was constructed *autonomously* and *systematically* after *questioning* and *analysing* that which seems to be self-evident. However, the elaboration of philosophical questions must be distinguished from the analytical abilities to perform thematic elaborations. Measuring with standards of correctness or maturity, Rosa may give evidence of a naïve, immature stage in the process of developing the concept of a horse.

This chapter describes thinking patterns in philosophising children. Section 3.1 describes and comments on the philosophical themes to be discussed, the development of philosophical notions, and a first order analysis of the progress of children's thinking patterns in their exploration of the philosophical themes. Children's procedural explorations are inspected conventionally and non-conventionally in Section 3.2. Conventional measurement begins with norms of maturity or correctness, evaluative criteria derived from outside or *a priori* norms. These procedural explorations can also be approached non-conventionally without any determined or well-known standard of comparison. Current (western and adult) standards may not be taken for granted. Without presupposed norms or

standards, the thinking patterns of children display a wide variety of alternative routes to find answers to initial philosophical questions, and to reach relevant philosophical notions. Section 3.3 examines some of the differences between philosophy and psychology, and the ways in which the disciplines have traditionally displayed themselves. On the one hand, the non-conventional approach may be most appropriate for detecting elusive potential philosophical qualities, like discovering ambiguities. On the other hand, situations leading to the performance of corresponding thinking patterns should be standardised to make patterns of philosophical quality comparable, and to comprehend a wide variety of performed concepts. To cover both approaches, thinking patterns will be empirically investigated in as open a manner as possible, while conforming to the three main philosophical features, and with the aim of laying a hand on these qualities. Thinking patterns performed by children will be observed as actual containers of theoretically based philosophical indicators.

All discussions recorded and presented in this chapter are among 11 to 12-year old children (from the last class of the primary school). They offer the reader an adequate and clear-cut thematic impression. However, these discussions have not been standardised according to scientific test procedures. Some date from a period prior to this study; others were used for publication ends and could not be stored in the databases used for this study. The discussions are not reproduced integrally. Recurrences and irrelevant remarks on illustrated themes have been omitted. Each description of such a discussion will be followed first by a short inspection of the philosophical content, themes and notions; secondly by observations with respect to children's explorations of the theme; and lastly by children's attempts to find answers on the initial philosophical question.

1 *Thinking patterns of children: samples*

This section examines six samples of philosophical discussions among children and aims to: a) acquaint the reader with the object of this novel approach; b) demonstrate some philosophical themes and developing philosophically significant notions; and c) show the procedures used by children to perform their thinking patterns. The original Dutch versions are presented in Appendix I. These samples address the following concrete questions:

1. What might be the start of the river Danube?
2. Is nothing perhaps something after all?
3. Can the countries on earth change?
4. Do endless roads exist?
5. What does voting incorrectly mean?
6. What is the truth and beauty of a poem?

Each sample will be followed by philosophical remarks concerning content and procedures that uncover a limited part of the broad range of possible thematic elaborations.

Start of the Danube

Steve: I don't think we can know where the Danube starts. All we can do is look at the patterns of the water, or of the ground; see which is the oldest.
Inez: Maybe one of the tributaries is the real Danube.
Prisca: Why is the Breche not called the Mindel, or the Mindel not Breche, and why isn't the Breche called the Danube?
Rein: That's what people have called it. But maybe you have to give it some thought.
Steve: The Danube might well be older than people.
Prisca: How do you know?
Steve: That's what you read in books.
Inez: But we still don't know where the Danube starts. If you start thinking about it, your thoughts aren't necessarily always right, are they? You can try to get it right, you can actually think: I'm right.
Steve: The truth is just a guess. And even if that guess is true, it was just a guess.
Inez: OK. One tributary is the longest one. That one can collect the most rain, and I think that it's the oldest one too. Well, it's possible. Because the oldest one forms the longest route, because it's had the most time to do it. Can't we have a look to see which tributary flows most like the Danube? About the oldest water: if it exists, it's been in the Black Sea for ages, and you'll never find out which river it comes from.
Jacob: Of course, it might have been rain?
Inez: Droplets of water in the Black Sea might have evaporated again and now in the Atlantic Ocean ...
Jacob: All that water must have been made out of rain. And if water is water ... Water has simply rained, hasn't it? Well, in that case it must surely have rained for a very long time? In that case, it's all just the same water, isn't it? I think that everything is the Danube. But where does water come from?

Philosophical content

Perhaps the most striking remark in this fragment is Steve's statement about truth. It carries a highly sceptical quality. Steve's reasoning seems to be a circular argument, but it is rather a subtle paradox, similar to that of the Cretan liar. It witnesses Steve's awareness of the arbitrary character of truth like that of Nietzsche. Another example of the arbitrary character of sentences can be noticed in the utterances of Prisca and Rein when they asked for names of the relevant streams. 'That's what people have called it. But maybe ...' They seem to say that names are just constructions to make the relationship between words and reality fit. Moreover, the fragment demonstrates several expressions of 'not knowing,' especially by Steve (first utterance) and Inez (utterances starting with 'But we still don't know ...' and 'OK ...'). Trying to get a handle on the identity problem, Inez reasons from space into time since the longest route offers the most opportunity to contain the oldest water. In her opinion, the river's identity depends on its shape, moving away from the initial question of where the Danube starts. Identity

is one of the main problems in modern metaphysics. When is it possible to speak about the 'same thing'? Were does the concept of an individual thing end? A well-known example of a comparable problem is 'the Theseus ship'; the ship that was saved and repaired for centuries, just to remember Theseus. But one can also easily imagine that the ship was built to carry on war and was damaged heavily each summer then repaired in winter so that every few years each piece of wood, each mast, nail, oar, or sail of the ship is replaced, so no authentic material remains. What is actually meant by the Theseus ship? Academic philosophers have many conceptual solutions for this problem based on concepts about the preservation of time, space, causation, or intention, and presupposition on complicated notions. In this discussion a similar kind of theme is at stake, although the Danube is a less definite object than the Theseus ship. Other themes concern the arbitrariness of naming and the delineation of the Danube.

Children's answer-finding procedures

The quest for the origin of the Danube seems to be driven by knowledge displayed at an early stage. According to Steve (in the first utterance), observations of the senses are the maximum base of knowledge; while Inez is likely to follow a more rational path of reasoning and deduction. Also remarkable is her trial to be right by genuine thinking 'I am right'. By this statement she imposes her intention on the relationship between word and reality. Questioning is, and will always be the starting point of philosophising. In this fragment, children seem to wonder almost continuously. 'Why is the Breche not called the Mindel? Questions and expressions of wonderment, including Jacob's last utterance, demonstrate a form of openness to new experiences and new knowledge. Finally, the frequent expression of tentative behaviour should be mentioned. Children seem to experiment continuously with cognitive problems and possible solutions. Their trials are often reflected in the use of special words like maybe and other expressions of modalities: *might be*; *you can try*; *it's possible*; *might have*; *can't we have a look?*

Nothing. Is nothing perhaps something after all?

Jochem: 'Nothing' is invisible. If it's in a bottle, you don't see it.
Casper: You see that there's nothing in it. So you see something.
Rémy: That makes 'nothing' on the one hand not anything, but on the other hand it's something.
Because we're talking about something ...
Timo: No! Nothing is not anything.
Berend: Maybe your eyes are deceiving you. Maybe there is something there, but you don't see it.
Rémy: But if you can't see it, how do you know that it exists?
Casper: Yes! Maybe there's actually a church over *there*.
Berend: But if your eyes are deceiving you, you can always bump into that church anyway.

Timo: There can't be a church there. A church is made by people and so they have to see it.

Wouter: You can have a sort of sixth sense effect, like: 'I see those people', – it's nothing, but you see them anyway.

Berend: How can you know if the things that you see, but you say that they don't exist, really don't exist?

Wouter: Well, because no one else sees them.

Jochem: Nothing can also be something that you don't know about. Like, if people say 'astromolieteblub', and you think, 'I know what that is: it's nothing at all'. And besides, 'nothing' can mean something, like when you find 'nothing' on the flea market.

Wouter: You see, if you taste 'nothing', you do taste something, namely 'nothing'. 'Nothing' on the flea market might be given a name. Then everyone will suddenly start selling 'nothing'.

Casper: But you don't say: 'wow, that "nothing" is brilliant'.

Wouter: 'Nothing' is what we can't name.

Marit: But with 'nothing' you notice nothing. You won't find 'nothing' anywhere. So the question doesn't make sense. What is the difference between nothing and something? Nothing? There's always something. Who can tell me what 'nothing' is and what 'something' is?

Timo: If something has a name, then I don't think that it's nothing. There was a time when we hadn't discovered the stars; we'd never seen them before either. Does that mean that stars don't exist?

Jochem: But if there's something that you haven't seen but it has a name, you can't say that it's something. You're not sure.

Berend: If researchers discover a new species of animal, they take a photo of it and show it to everyone. But at that point the species still hasn't been given a name.

Wouter: If a baby has never seen or smelled pickled cabbage, it's still pickled cabbage, isn't it?

Timo: But you don't know that at the time.

Marit: But what's in a bottle that contains 'nothing'? Invisible pickled cabbage? There has to be something there.

Jochem: Maybe 'nothing' is a group of objects that we don't know much about, and that are therefore called 'nothing'. On a flea market you see vases, a trumpet, old things ... But you're not looking for them. We know the word 'nothing', and it does mean something, but we can't see it. We've never seen it either. This is really a very tricky topic and it's very weird.

Philosophical content

The quest for *nothing* is sophisticated because something is named that does not exist. No object exists that fulfils the concept. The elaboration of the concept of *nothing* brings to mind the Chinese philosopher Wang Fuzhi (1619 - 1692) who stressed that *nothing* is nothing (van der Leeuw, 1994, p. 319 - 320). In contrast, Heidegger and Sartre stress some quality of *nothing*. The discussion deals at the same time with the question of whether nothing functions as a name or as another

logical syntactic entity. Non-existence may also be interpreted as ambiguous. Already in the first sentence, the concept of *nothing* is used with two different meanings: as object and (dis)-affirming adverb (i.e., considered as negation of seeing). A comparable problem occurs in other sentences, as with Wouter's comment about the tasting of *nothing*, and Jochem's emphasising observing *nothing* on a flea market. *Nothing* is interchanged between the category of objects and that of (dis)-affirming adverbs. This might be done unconsciously, and interpreted as mistaken, or consciously and interpreted as a game, as in many jokes and riddles. In the Netherlands jokes about Belgians are popular. One may ask for example: 'Why do Belgians put empty bottles in their fridge?' The answer, of course, is that this must be to provide for someone who likes to drink nothing! Jokes and riddles may be considered as fun and as nonsense, but they may also expose a play with categories or classification criteria. They challenge truth or the correspondence between thoughts and reality, and exploit vagueness and ambiguity. Children exploit the pragmatic meaning of a word and require semantic, syntactic or semiotic meaning references to fit it. Knowing and experiencing seem to move into one another seamlessly. This reminds one of the tendency to pragmatism that can easily be detected in the thinking patterns of children.

Children's answer-finding procedures

Trials, errors, and games are probably strategies and non-conscious routes for understanding the *nothing* problem, and to serve master concepts. Children need to test and to play with extensions, classifications, and the conditions to apply terms. Procedures of testing and playing may be judged as mistakes, as reflecting a development, or as games with open-ended unconventional thinking patterns. In any case, this judgement depends on cultural backgrounds, scientific presuppositions and biases of the interpreter.

Children ask many questions, and many question marks can be noticed in the discussions. Children are easily surprised and like to explore this. They ask about the uncertainty of knowledge; they suppose that senses may deceive us; they try to detect relationships between knowing and observing, proving and naming ("nothing is what we can't name"). They are puzzled. They use the word 'weird' in the last sentence, not to disqualify the problem as many adults do, but to identify it just as 'really puzzling'. Problems seem to be problems by their quality of being weird. Trials can also be recognised in expressions of modality, in the use of words like 'sort of' or simply by suggestions. Casper exposes his tentative behaviour in a quite pushing suggestion *nothing is brilliant*. Also highly imaginative fabrications like that of *invisible pickled cabbage* and fantasies about a *sixth sense* as a possible solution can be interpreted as tentative behaviour.

The reasoning patterns exposed are also remarkable. The children exploit the ambiguity of *nothing* and rearrange the conditions of applying *nothing* arbitrarily: How can you know that nothing really doesn't exist? "Well, because no one else sees them." (Wouter). They explore the linguistic character of reality and try new words (*astromolietblub*). They search for rigid logic rules to demarcate nothing

and something. Sometimes, they follow different trains of thought; sometimes, their arguing does not seem to be proper in every detail, and leaves the possibility open to interchange divisions of the arguments. So, for example, does the following reasoning: When Berend states that "deceiving eyes may lead you to bumping into the church", Timo went on to prove the existence of the church by stressing that it was constructed by men. This explanation seems to be based on 'division and composition' of relevant and irrelevant arguments.

A final characteristic of thinking patterns in this fragment is the way children employ illustrations or anecdotal sediments of 'nothing' in several everyday experiences. This way of expressing thinking patterns can be identified as a smooth transition from abstract into concrete; or as a trial to discover different meanings. A smooth transition from abstract into concrete is performed by associating, and by evoking images of, for instance, the flea market and a bottle containing invisible cabbage. Illustrations and anecdotal qualities may be significant in a jointly performed trial to understand the problem of *nothing*. This transition or trial can also be seen as an investigation into the effects of *nothing* by varying syntactic and semantic contexts.

Can the countries on earth change?

Rex: Countries can and can't change. Countries can change on maps, but not in reality.
The ground can change. After a volcanic eruption, lava can turn into new ground. And technology just keeps on developing. That's why we now think that the world is round instead of flat. So maybe there'll be more new insights in the future.

Najia: You can make Morocco bigger too. If everyone wants Morocco to be bigger than the Netherlands – maybe because a lot of Moroccans live in Germany, in the Netherlands, in ... Well, I don't know, everywhere, who want to go back to Morocco.

Lisa: But how can you make Morocco bigger? More houses ... OK, but that doesn't make Morocco bigger, does it?

Willem: Well, using sand... but if you get sand from somewhere else, then *that* place will have less sand.

Esther: Yes, I mean, where would you get the sand from? From the Sahara Desert? But that would make the desert smaller.

Ramona: The Sahara wouldn't get smaller, only that particular sand dune. You lose sand, but that still doesn't make the Sahara smaller.

Rex: In the future, there might be a machine that can defy gravity. And then you could, like, remove a country and put it onto Morocco. But then you'd have another problem: that bit of land won't be Morocco anymore.

Tjalling: But, for example, if Morocco went to war with Algeria, and if it wins, it will get more land.

Ramona: Well ... Najia said that the people would have to want it to happen, but I don't think that the Algerians would want that.

Lisa: You can move the world by going to war. If two countries, like Spain and Italy, start a war against each other, and Italy wins a bit of Spain, Spain a bit of Italy, and this goes on and on, you get to the point when Italy has won Spain and Spain Italy. Then Italy would be in Spain and ...

Rex: But look, Lisa, you're saying that they've swapped, but it's only really the name. I can call the Netherlands, Mars, but that doesn't mean that the Netherlands *is* Mars.

Lisa: But the Netherlands has become the Netherlands, and Italy used to be something else too, didn't it? A country can become something else now too, can't it?

Rex: Yes, that's what we think, but maybe it's really different.

Tjalling: You can make one country bigger with wars, but the other one will get smaller.

Rex: We *can* make things bigger actually. Because if you put an ant, for example, under a microscope, it will look much bigger to you. Then all you need to know is the formula for how to, like make it bigger for real.

Philosophical content

In this fragment, the most striking and finest example of children's thinking patterns is undoubtedly presented in the last sentence. By referring to an ant under a microscope, the difference between 'bigger in the *eye*' and 'bigger for real' is presented as a matter of a formula. This emphasises an underlying relationship between the observable world and reality. Rex seems to be aware of a rather sophisticated theory of truth. He is playing out the consequences of his train of thought about *making bigger*, manipulating the correspondence theory of truth, with a rigid logic. This theory, displayed for example by Aristotle, claims a correspondence between knowledge and reality. He is experimenting or playing with the conditions of applying terms and with the assignation of extensions to concepts. The arbitrary character of reality was exposed earlier when the same child stressed the name of a land representing its identity. He continues: "I can call the Netherlands, 'Mars', but that doesn't mean that the Netherlands *is* Mars". So, Lisa's tentative reasoning on going to war must be interpreted in the same perspective. The entire discussion seems a major exploration into the criteria for identification of observable objects (countries) and for the reference of their names, for instance through conventions of agreements. At the same time, children are also exploring potential meanings of *changing* and *becoming bigger*.

Children's answer-finding procedures

In addition, children appear to be able to reason rigidly and rationally – like playing a game – without any connection with sensory experience, leading to a rather tentative conclusion that criteria for naming and knowing reality are arbitrary. Wonderment, puzzlement and the sensitivity to ambiguity seem to be embodied in often used words like *actually* and *really*. This sample also demonstrates tentative behaviour and the use of thought experiment to elucidate philosophical questions.

The discussion exposes trials par excellence through *for example*, *new insights*, and in the suggestion *there might be a machine that can defy gravity*. Najia is even trying to impose her intention on reality to change lands. By thought experiments about ground, sand, the conservation of mass, and inhabitants, children search jointly for conditions of applying *land* and for assignation of extensions to *land* and to expansions (*making bigger*) in respect to lands.

Endlessness

(On the possibility of endless going further)

Andrea: In a train you can keep on travelling, past towns, and stuff. You see more and more tracks all the time. It doesn't stop. When does endlessness actually stop?

Ron: There's always an end. Each train arrives at a terminus.

Paco: But the rails just keep on going.

Jerry: With motorways you've got the A2, the A9 and ... If you carry on it turns into the A3, for example. Well then, the end of the A2 turns into the A3.

Paco: Then that road just gets a different name.

Jerry: Yes, so it's a different road. Because, look, on the A2 you're allowed to drive up to 120 km but only 100 km on the A3. That makes it a different road, doesn't it?

Umi: If the A2, A3 etc lead in the same direction, then you can drive on endlessly, can't you? But if you come to a bend, that's the end of the straight road.

Jerry: Yes, take the Huygenstraat Street – it's covered in road clinkers. That street turns into the Tesselschadeplein Square. That's covered in tarmac. It's a different road surface and therefore a different street.

Andrea: Endlessness is when you keep on going. There's no end to it. That applies to driving around in circles too. Anyway, I think that endlessness exists. If a road is given a different name, that doesn't automatically make it a different road. Motorways are just *one* network of stretches of tarmac that lead all over the world.

Paco: Maybe you can travel to the vanishing point, with enough power, oxygen and maybe you're actually immortal. Maybe you emerge somewhere where you don't know anything. So, maybe endlessness does exist, but maybe not, because you don't know it at that point in time.

Ferhat: It's an illusion. Why do you call it endless? Call it endlessly.

Philosophical content

The idea that the identity of a road is carried by its name, and thus by language, is very clear in this fragment. The relationship between the concept of a road and reality is explored again by their names and rules attached to those names. A road labelled as A2, determines the road's identity together with its rules. So the description of the road overlaps with its prescription, as embodied in a set of rules.

This discussion displays several criteria used to identify the object of a road and offers an opportunity to discern differences between logical, conventional and practical consequences of using the concept *road*. Although observable properties of concepts are relevant in general; in this example, the road's observable and textural appearance seems to be only a question of something that has to be mentioned in passing. Before, but also after stressing these properties, children search other perspectives like those of direction, movement and final goals. This leads simultaneously and in addition to the issue of naming and its criteria, to a search into boundaries and infinity (comparable with the discussion on the Danube). While the various types of boundaries (demarcations) are prescribed by names, rules and appearance, the concept of infinity addresses a theory of endlessness and an idea of everlastingness. Endlessness is explained by stressing the meaning of *going on*, with its temporal and spatial components. By analysing the practice of endlessness going on forever, Andrea moves to circles and to a network. Relativity is introduced. In Paco's next sentence the smooth transition from endlessness in space into endlessness in time is elaborated into immortality.

Although the starting question in this discussion referred to the road's endlessness, many other philosophical themes were explored. It is the challenge of every philosophical discussion to find and maintain the focus of the discussion, starting with its thematic question. But it is also the challenge to find a middle course between this theme and wild thoughts and associations that result from the thinking process generated by the initial key question. For this reason, it is difficult to classify philosophical discussions into separate theme oriented groups. This point will be taken up again in Chapter 5.

Children's answer-finding procedures

Through Paco's very speculative utterance, the conventional meaning of an endless road seems to be stretched or to slip into figurative meanings. This trial may lead into vagueness, but it also creates an opportunity for a wider concept expansion. Vagueness is part and parcel of meaning (Bartsch, 1987). Through vagueness, concepts are not constrained by correctness, so that they can evolve fundamentally to an unlimited extent. By the multiplicity of interpretation levels of the concept *road*, thinking about the identity of a road seems to challenge the thinking about space and time in general, as through metonymic and metaphoric use of the concept. However, this does not imply that children are conscious of this wide meaning framework.

Many sentences end with a question, and seem to represent attempts at answering. They are expressions of assaying or tentative behaviour. This is emphasised by the frequent use of *maybe*. The fragment also exposes a high degree of creativity with its many suggestions, fantasising thoughts, and attempting a solution for the emerging problem by trying a new word: You just *call it endlessly* (last sentence of Ferhat). Finally, in Paco's *maybe yes but maybe not*, he seems to express his sensitivity to ambiguity.

Voting incorrectly

Marlon: I think it's strange that someone who can't think clearly, can still vote. He might vote wrongly!

Desi: Someone who can't think clearly is a person too, isn't he? He's not an alien, is he?

Nina: Everyone above the age of 18 is allowed to vote, no matter whether they can think clearly or not.

Jos: Suppose he had to vote for 'the world must be wiped out immediately' or for 'the world must not be wiped out', and he presses the wrong button ...

Nina: But which is the wrong button?

Lucy: You can explain it to him, can't you?

Marlon: He might choose something wrong, something that he doesn't want.

Lucy: It's never wrong. Maybe he doesn't know what he's doing, but wrong ...

Nina: The button he presses is *his* choice, *his* thought.

Jos: Other people, who can think clearly, they can have thoughts that aren't good too.

Lucy: It's still his opinion.

Marlon: Of course not. He doesn't even know what he's voting for. He doesn't know what his opinion is either.

Desi: Maybe his mother can say: just press that button, because I think it's the right one.

Marlon: Does his vote reflect *his* opinion in that case? No way! He's just pressing a button because he thinks that it looks nice. And what if he chooses the wrong one?

Lucy: There is no wrong one. Because you'll never have to vote for or against: should we end life on Earth.

Nina: Wrong and right do exist, but not in voting. There are just parties that want to win, and there are never any important issues to vote on such as, 'should we let life on Earth die out?'

Lucy: But look, if you say that some people can't vote, you can just as well say that some people can't go to the funfair.

Jos: Yes, that's a good one; because, like, someone doesn't think that he might get really sick on the rollercoaster.

Nina: If he goes on it anyway, it's his own choice, isn't it? It doesn't kill you, does it?

Jos: No, voting wrongly doesn't either.

Nina: Right and wrong are different in this example.

Desi: No, it's the same.

Marlon: But voting does have something to do with thinking things through, and not with what button you think looks best.

Lucy: Professors can vote wrongly too. Everyone can make mistakes.

Desi: My grandma is always forgetting everything too.

Nina: Really it's stupid to say that only over 18s are allowed. When you're 10, you might be able to think things through clearly too.

Jos: You're the only person to know if you can really.

Marlon: Yes, which button you think looks best. Maybe you vote correctly. Maybe you think that voting correctly depends on watching the news on TV.

> Well, that guy in the 'Goede Tijden' series is 38, but behaves like a 6-year-old. If a 38-year-old is hit on the head really hard and behaves like a 6-year-old, is he allowed to vote? I think it's weird. Does that mean that you seem 18 to yourself or to other people? Is it a question of thoughts, or your body?

Philosophical content

The most salient philosophical feature of this discussion is the entanglement between wrong and incorrect. The Dutch words *goed* and *verkeerd* are used and translated into different words for each notion: *clear/right* and *incorrect* refer to logical meanings. *Good* and *wrong* on the other hand, refer mainly to moral meanings. This confusion is shown here as a mistake, but is also questioned as a serious point of enquiry. Several categories of (in)correct voting are interchanged: voting in concordance or discordance with moral good, voting with or without opinion, and voting in concordance or discordance with someone's intention. This interchanging can be considered as a mistaken act of classification. Ambiguity is used or exploited unconsciously here. On the other hand, children themselves are searching actively and consciously for the division in meaning of the concept *good* and in that of the concept *wrong*. Nina's question – *But what is the wrong button?* – is not exposed here as a riddle, joke or game, but as a trial. Besides possible literal meanings, *the wrong button* may be explained metaphorically. Searching for the meaning of *right and wrong*, children try out different uses of words in different situations or contexts. *Wrong* or *incorrect* can be applied to the object of voting: political party, to thoughts, to the succession of thoughts, to choosing a button. But it may also refer to *not knowing* or *making a mistake*. In the thinking patterns displayed, *wrong* (in Dutch: *verkeerd*) is mainly assigned to cognitive verbs: choosing, wanting, knowing, thinking and judging (nice). In particular, the combination with a cognitive term makes *wrong* susceptible to moral interpretations. *Wrong*, in combination with cognitive verbs, can be interpreted logically and morally. Although many children cannot explain this contrast, they do feel uneasy with it. Something with respect to the meaning of *wrong* lurks. As a consequence they don't stop asking the same question and continue outlining the same point in different illustrations, or in different words. They perform such patterns in order to obtain clarity finally. Another philosophical theme that is explored is the anthropological question. The human being is sensed through contrasting him with an *alien*, or analysing and comparing different human qualities like in the last sentence: *is it a question of thought or your body?*

Children's answer-finding procedures

Besides the ambiguity of right and wrong, analysis of reasoning patterns two other forms are shown: reasoning through dichotomy and reasoning through analogy sometimes based on associative links. Dichotomy is illustrated in Desi's sentence, when she makes the quality of human subject to yes or no to being an alien. Henceforth, Lucy and Nina argue the non-existence of right and wrong by stating

the impracticability of two phrasing questions. By the exclusion of the dichotomy of positions in case of elections, they come to the conclusion that *wrong* with respect to voting is not possible. When Nina is arguing the right or wrong of voting by emphasising that *it doesn't kill you, does it?* she points to an inadequate consequence and falls into the same 'fallacy'. This dichotomy reasoning is aimed towards the discovery of demarcations strictly limiting the questioned concept.

On the other hand, analogical reasoning makes concepts vague and wider. Many analogies are possible, but what analogies are considered to be acceptable? Comparison of the voting problem (i.e., being allowed to vote) with being allowed to go to the funfair (Lucy) opens the mind for presuppositions and makes the problem indefinite. Comparison of being allowed to vote between persons at different levels of clear thinking, like that of a professor and someone behaving like a 6 year-old after being hit on the head, does the same. A presupposed clear relationship between the capability to think clearly and capability to vote is under review again. Children come to conclusions like: *Right and wrong are different in this example* and *Really, it's stupid to say that*, stressing ambiguity and vagueness of the concepts. In a discussion on a similar theme, described in 'Jong & Wijs' (Rondhuis, 2001, p. 69), children compared age limit in elections with age limit in entrance and admission prices of funfairs. They concluded with the arbitrariness of age limits. Even more striking in the same discussion, was the conclusive sceptical stance of memorising that a single good answer cannot exist because there will always be right and wrong answers in different languages. Since there are multiple correct answers, what constitutes a correct vote?

Children display tentative behaviour through use of *supposing*, modes of modality (*maybe*, you *might have*), questions for verification, and by suggesting other interpretation schemes. The start and finish of this fragment show expressions of wonderment through words like *strange* and *weird*.

Valuing poems

A final sample is taken from a discussion that appeared in a Dutch newspaper (Trouw, October 2001) carrying poetry as a leading theme as part of Children's Book Week. Children from 11 to 12 years in age (group 8 of the primary school) were asked to philosophise on selected poems, searching for their truth and beauty. They reasoned about correctness, clarity, lucidity and lunacy of several poems, in particular one by the Russian poet Alexander Blok (1912).

Night, a street, a lamp, a pharmacy, a useless and dimm streak of light.
Even if you live another quarter of a century,
Everything will remain the same. There is no way out.

You die and you start once again,
And everything repeats itself, as of yore:
Night, the frozen ripples of the canal,
A pharmacy, a street, a lamp.

(Translation: Gerard Brugman)

Umi, a girl of 12 years old, argued against a classmate who was disgusted by this poem.

Umi: Well ... still ... the poem is a little bit about 'what sense has life?' Life just keeps going on. Everything is connected with each other. Sooner or later the pharmacy comes to the streetlamp. In the pharmacy you have light and the lamp gives you the light. The poem makes you think: 'Would it be true?', 'Can somebody become something else?' Some poems are not in rhyme, but still they are poems, because they give you a sort of 'feeling' (in Dutch a verb, *indenken* is used): a thought that can also happen to you. You can wish for something weird. One time I wanted to be a flea. What would it feel like? It sounds so logical, but it would be all different: only big things around you, you don't go to school, real strange. This could be a poem.

Philosophical content

Aesthetics is a significant philosophical topic that is generally missed by children. Without trying to look at the deeper artistic denotations and relationship between form and content, children's judgment of artistic expressions seldom exceeds their human dimensions. Topics like *'When is a monkey's painting a piece of art?'* and *'Is art in Amsterdam the same as art in Timbuktu?'* or *'Is beauty nowadays the same as beauty in the Middle Ages?'* are mostly interpreted by children in their anthropological, social and ethical aspects. Nevertheless, this poetry discussion still leads to questioning and wonderment about the free transition from one domain of experience into another, and stresses the intellectual distance between performer and performed content in the act of thinking. The free transition between domains of experience can be considered as a poetic capacity and interpreted in linguistic terms dealing with metaphors. Here, a metaphor of a geographic environment is used to identify circularities and slipping time in different forms of existence, in dressed streets and in the narrative of a life. The transfer from thinking to feeling is emphasised by way of poetic expressions. The intellectual distance between different positions of experiencing the same event becomes an explicit philosophical content in Umi's vivid description of the distinction between being a self and being something else, for example a flea.

Umi's answer-finding procedures

It appears that Umi has a well-developed sensitivity to poetic qualities, sensing the disunity between a real life version of events and a magical interpretation of these. She senses the truth of the circularity of things, as with the streetlamp and the pharmacy at night, and of life in the repetition of dying and starting again: Escaping or changing does not seem possible. She does so without clear marks of consciousness, no more than understanding intuitively without apparent signs of reproduction. Her justification is rather naïve, emphasising the need for the streetlamp when it is dark in the pharmacy. But she keeps questioning herself, attempting and trying out new solutions, for example by inventing new word

uses (Dutch: *indenken*). This term detects thoughts that take readers by surprise, making them part of a solution and generating other thoughts. Umi seems to make conscious use of regressive abstraction (Nelson, 1970), looking at conceptual structures lying at the root of her experience. She explains her feeling of a similar 'something' on two different occasions. She compares her imagination (thinking) of the detected circles in the poem with her imagination of being a flea. In both cases, she is wondering and puzzling over the truth and about the realisation of this imagination. Knowing whether it is true is a form of speculation. In any case it is *weird* or *strange*. The thought of asking herself how it would be being a flea is directly linked to role taking and demonstrates an intellectual distance between her and her thoughts. Umi's approach to the presented poem leaves the impression that the detection of philosophically significant notions are within reach without being able to justify the events.

Conclusion of these demonstrations

The discussions presented demonstrate how philosophical themes and notions are elaborated upon in children's thinking patterns. To determine the philosophical quality of these patterns, certain combinations of characteristics seem to be significant: 1) tendency of searching for truth; 2) often combined with a sceptical attitude of not knowing on a meta level; 3) wonderment and puzzlement; 4) tentative behaviour; 5) ability to identify the syntax of concepts, identity and meaning with its criteria. Part of this ability is the discovering of relationships, like the relationship between language and reality, ending with the detection of its arbitrary character and the relationship between intention and observable world. However, these findings are crude generalisations centred on themes like origin or identity. In any case, the children are thinking and discussing in response to a problematic question in order to reach clarity. Questions are apt to generate ambiguities, vagueness and uncertainties because of the non-existence of definite answers. A final answer to identify the nothing, the start of the Danube, the invariability of countries, the end of roads, the correctness and good of voting, and the correctness and beauty of poems, seems to be inaccessible. The initial, key questions were identified as philosophical in advance and lead to discussion of philosophically significant themes and notions. The highlighting of philosophically significant notions can be considered an automatic result of thinking about philosophical questions. What remains to be judged are the procedures required to arrive at these outcomes.

Through the procedures employed to derive answers, children explore philosophical themes and notions by sensing them, especially their ambiguity, vagueness and uncertainty. They analyse and reason concepts, relationships, connotations and associations. They explore boundaries and different interpretations. They search for smooth transfers from abstract thinking levels into practical occurrences, and vice versa, and from one domain of knowledge and experience into another. These explorative thinking patterns are philosophically qualified because the main philosophical tendencies can be identified in them. Sometimes the expressed thinking patterns will be legitimized through current standards of logic, by

requirements of correctness, norms of maturity, or presupposed ideas about usefulness. But since part of the main features of philosophy is rooted in a character of non-determinability, efforts to measure the philosophical qualities of exchanged thinking patterns described in the samples are best served by approaches that go beyond the identified criteria. The following sections will focus on the different approaches to study thinking patterns in children and reflect upon their answer finding procedures to philosophical questions.

2 Evaluative approaches of children's answer finding procedures

An understanding of the procedures children use to search for answers and express their thinking patterns requires 'something' to verify. Typically, verification is the application of criteria (norms) in an evaluative approach: a model in which observations and thoughts can be structured. The actual appearance of thinking patterns can be tested to see if they fit such models. The criteria or norms vary widely. They refer to: 1) fundamental paradigms dependent of times, theory and culture; 2) requirements of correctness derived from specialised disciplines (logic, language, mathematics); 3) norms of maturity and other performances which can be assessed empirically; 4) cultural justified 'agreements' or institutional consensus. Consistency is an example of a logical requirement. Ideas about usefulness and political correctness are examples of institutional consensus based on more or less well-reasoned presuppositions. Consequently, thinking patterns will be interpreted as intelligible or obscure, as useful or absurd, as correct or mistaken, as stage-developed, or immature. Sometimes children perform correctly. Often, their performance is judged as 'mistaken' on account of their immaturity. To judge thinking patterns beyond omnipresent norms originating in educated western styles of reasoning, particularly in the performance of science, will require a significant shift in traditional norms.

Western model of thinking

It is difficult to think beyond our mature western educated model since it is used in almost every judgement. This model seems to be part and parcel of our thinking patterns, supporting the hypothesis that growing to maturity involves thinking problems that can be adequately tackled in a socially acceptable way. This excludes the possibility that children may develop better thinking patterns than adults. Although such ideas may lead to theories concerning gradual development into mature performance and recapitulation views (recognising primitive thinking patterns in the way children reason), no linear development of philosophical thinking patterns can be detected in history. Different traditions of philosophising depart from different orientations and show a different succession of ideas. In the western tradition for example, the world is expressed in terms of distinguished objects, situations and events. In ancient Greek thinking, world or reality was considered to be a collection of *substances*. Thinking about development and the need to explain the occurrence of changes are inventions of the western form

of thinking. In the eastern thinking tradition, reality is a collection of *processes*. Where Greek philosophy has to compromise in order to clarify changes and development, eastern philosophy does not need to take pains to account for this phenomenon since changes are naturally incorporated by their conception of a reality oriented towards a practice of acting (China) or towards liberation of earthly things (India) (Bor & van der Leeuw, 2003).

West and East developed different philosophical notions, for instance with respect to ethics. While western ethical thinking is guided by more or less objectively acquired criteria related to different notions and concepts outside the individual, in the east, ethical criteria are supposed to emerge from the individual. A clear criterion for an outsider to judge the moral value of an act, is missing. This difference has significant consequences for judging moral development, particularly staged development. One may easily argue against considering individual development as a recapitulation of historical developments (Koops, 2001). Incidentally, similar notions or thought constructs recur in history. Academic philosophical approaches to human contact with animals may demonstrate this phenomenon. In modern times, these contacts are based on the idea of creatures with 'equal' standing. In prehistory, or in so-called 'primitive societies', people dealt or deal with similar ideas. Likewise, some stages of development can be identified in inverse succession, and some appeared to be skipped altogether. Obviously, no necessary succession of arbitrarily formulated stages in the development of thinking patterns exists.

Immaturity

It is difficult for adults to think beyond ideas of maturity when judging children's performances. Deviation from what is held as a sign of maturity is easily interpreted as immature. Children often use concepts linguistically with extensions other than those accepted by adults, such as those of time and space. In conventional developmental psychological terms: they are used with incorrect extensions. This is especially true when a concept carries a vague extension, such as *horizon*. Once a twelve year old boy described a horizon as 'an invisible line in the distance that is going farther and farther.' However, he did not mean that the line could not be seen. Upon being questioned, he added the following sentence: 'you are able to see it, but you are not able to reach it.' Apparently he was using 'invisible' in connection with 'going further and further' to emphasise the extraordinary character of *horizon*. A horizon is not in the distance like stars: stars are within reach and a horizon is absolutely inaccessible (Rondhuis, 1994). Because children are supposed to have less experience than adults and incomplete views of the world, they tend to think in fragments. They miss the ability to explain as much as adults or seem less skilled in using socially acceptable explanations. Many children overcome this inconvenience by bridging experience gaps and developing non-conventional relationships. Children use linguistic and reasoning styles as instruments to fit things into their thinking frames and into their conception of reality. This is done by stretching these frames through the ambiguous use of conjunctions. Conjunction like 'but', 'and', 'because', 'for', 'thus', 'so', are often

exploited with the sole meaning of 'including'. The original conception of the world has to be expanded by an additional event. In *The Philosophy of Childhood*, Matthews (1994) describes the wondering of his four years old daughter Kristin about colours. While painting, Kristin told her dad that her world is made of colours. When asked: 'What about glass?' she declared the world to be constructed of colours *and* glass. When faced with something that seems opposed to an achieved view, reality will be provided with an additional dimension.

Object and subject of studying

It is also important to stress that children are not only the objects of study; they are also fellow citizens with their own respectable thinking patterns. Adults should not condescendingly judge philosophical thinking patterns performed by children because they cannot easily be evaluated in adult terms encompassing adult's bias. A child's own evaluations may be equally valid. Children's thinking patterns have to be approached as open as possible. The measurement and formulation of criteria based on the supremacy of maturity would be premature or even presumptuous. On the other hand, one has to guard against judging childhood thinking patterns according some romantic view concerning children as heavenly creatures.

Requirements of correctness: linguistic performance

Linguistic performance is culturally regulated by linguistic norms of correctness and maturity. According to many psychological handbooks comprehension of language precedes the production of it (Berk, 1989). Conclusions about achievements concerning comprehension are difficult to derive from linguistic production, because mental states and processes cannot be directly observed. Meta-linguistic awareness or the ability to think about language demonstrates the distinction between linguistic competence and linguistic performance (Chomsky, 1957). A discrepancy between thinking and language abilities interferes with interpreting children's expressed thinking patterns. Concept formation is studied extensively by using several types of norms. A correspondence between mental concepts and their verbal expression is presupposed. In their exploration of philosophical questions, children search for similarities and contrasts between words and their various conventional uses, and also try out new words. Their linguistic performances take on the form of magical instruments used to overcome apparent antitheses and contradictions (Rondhuis, 1990). Familiar meanings are stretched. Discussions with children often contain examples of overstretched meanings.

In a dialogue with five 11 to 12 year-old children searching for the meaning of *promise*, they discussed the question: Is a promise everlasting? One boy continued: "When it is impossible to pay off the money of a loan on the promised date, you *can make the loan longer*." (The Dutch text: "... je kunt het *lenen* ook *langer maken*, zeg maar.") Although the boy did not use a word in Dutch that expresses a period or duration, he tried to expand the concept *loan* with a dimension of extension that it does not possess by using the expression *longer making*.

Another example is illustrated in a discussion about time and the length of a day when a girl (age 12) noticed: "Maybe a day is not ending when it becomes dark ... Then it is a day without an end: it is kind of *flashing day*: on, out, on, out ... There may be light, but not during the whole day. A flashing day is an everlasting day" (Rondhuis, 2001). Here, it is demonstrated that a correct concept of daytime does not actually exist. It also is a fine illustration of the course from linguistic performance into reasoning and into relativistic ideas about the character of time.

Requirements of correctness: analytical fallacies

Reasoning comprises the discovery and use of meaningful relationships between 'things' (i.e., objects, situations, events, or mental images). Meaningful relationships presuppose a norm of meaningfulness, referring to successful orientation in reality in a cognitive sense. However, the discovery of meaningful relationships may go by non-conventional lines or may be judged as mistakes using *a priori* criteria. Children may carry out meaningful relationships with a strained or artificial severity and demonstrate a straight, rigid logic in their reasoning, but then detect quite understandable trains of thought as outside that rigid logic. Having accepted a system of thoughts (and events), children may artificially force apparent non-compatible life experience to fit into this system. So for example, an 11 or 12-year old boy did not differentiate between *being a cat* – in reality, and *being a cat* – in play (i.e., behaving like a cat). The identity of *being a cat*, in reality and in play is true only if properly played according to a system of rules to perfectly simulate reality (e.g., by wearing a cat's skin). Then you would look and feel like a cat! If not, the performance fails (Rondhuis, 1990). Game, play or reality, this rigid logic only lasts for its duration. No differentiation between the system of reality and the system of play appear. Lines of thoughts are captured artificially in the rules of game or play, as in rigid systems of logic that end in the belief of the arbitrary character of reality. According to standard adult forms of judgement, this type of reasoning can run into an obvious mistake.

Often, fallacies can be interpreted as incorrect acts of classification. The discussions of the foregoing section demonstrate some interchanging or incorrect use of categories. The concept *animal*, as 11 to 12 year-old children generally understand it, offers example of interchanging categories. Many children pretend that eating meat is allowed, but that eating animals is forbidden. Here, animals are opposed to meat generating entities; and have nothing in common with respect to eating. Evidently, two categories exist to represent one 'adult' category: 'animals'. In children's terminology, animals are different in a biological sense as well: they need not to be alive. Pets are also considered as animals. In the eyes of these children, the possibility of experiencing a profound relationship with the animal is considered as the most essential property. Objects in the category of animals are identified and classified with different criteria when applied by children and adults. Sometimes literal meaning components are challenged and exaggerated. For example, the Dutch word for vermin, *ongedierte* (*non-animal*) is often interpreted as the opposite of animal. Reasoning logically, children come to the conclusion that vermin are not animals. The ability to evoke contradictions in

reasoning conforms to dialectic ideas. According to Riegel (1973), children tackle questions on their own level. 'Mistakes', as they are called in evaluative approaches, can be considered as the driving force of the dialectic process.

Presuppositions concerning morality and usefulness

Poorly developed presuppositions often appear in evaluative approaches to culturally justified requirements of political correctness and thoughts about good and evil. For example, many would question someone who prefers to buy six treacle waffles at a market place for €2.00, when ten waffles cost €3.00. Apparently, this conclusion is arrived at by certain presuppositions about a relationship between price and quantity. Of course, it is possible that the person buying the waffles could attempt to anticipate or prevent confrontation over remaining waffles and superfluous waste. Following the same line of thought, many 11 to 12-year old children are amused when told that prison cells are provided with televisions and other instruments of entertainment: Prisoners have to suffer so prisons should provide very simple facilities. One has to conclude that there are hidden presuppositions in many events that are valued in daily life. Traditionally, it is a philosopher's task to uncover them. Therefore, an evaluative model based on institutional consensus or poorly reasoned presuppositions is not appropriate to judge children's thinking patterns.

Evaluative approaches

Thinking patterns should be judged on their procedural merits in order to arrive at philosophically significant notions. Conventional measurements or evaluative approaches based on norms of maturity, requirements of correctness concerning language use and reasoning styles, or cultural justified 'agreements', detect fallacies and mistakes. Generally accepted condemnations and reasoning have their roots in hidden moral or economical presuppositions. At the same time, the procedures judged as 'mistaken' according to these evaluative approaches could be interpreted as a method to challenge concepts, to stretch and exaggerate meaning components, to classify entities in different categories, to try new words, to create unconventional relationships, and to tackle questions on an individual level. The sample discussions show a high degree of such tentative behaviour. 'Mistakes' do not seem to matter when detecting the general constituents of philosophical thinking patterns. Since the measuring methods of evaluative approaches are based on detecting mistakes, these approaches do not adequately measure philosophical qualities of the exchanged thinking patterns. The exploration and judgement of children's answer finding procedures expressed in their thinking patterns is best served by approaches that do not incorporate the conventional criteria of these evaluative approaches.

3 Approaching children's answer finding procedures beyond conventional norms

If conventional evaluative approaches fail to detect philosophical characteristics, how can thinking patterns be investigated? The observation and sensing of thinking patterns displayed in this chapter must uncover tendencies dealing with uncertainties and the autonomous production of thoughts on unanswerable philosophical quests. If presupposed norms of maturity are discussed, requirements of correctness will be paradoxical because their correctness cannot be determined. As a result, poorly reasoned use of cultural justified agreements would be identified as non-philosophical. What other criteria can be used?

In this section, initially, the focus will be on expressed thinking patterns, observing and investigating them as openly as possible. They appear as a collection of events, unstructured in advance. Secondly, an evaluation follows. For example, the thinking patterns of ten year-old Rosa, concerning the question: 'what is a horse?', may be evaluated as naïve, uncultivated or odd. However, this judgement was formulated without a well-determined philosophical touchstone or measurement: neither with respect to the horse, the object of thinking, nor to the act of thinking or to Rosa's intention. Ignoring presuppositions of maturity, requirements of correctness, and cultural justified agreements, a variety of alternative thinking procedures and views can be seen. Children explore the offered philosophical themes tentatively through: 1) reluctance to impose hierarchical relations between objects or events, reflected in the juxtaposition of them; 2) questioning, wonderment and puzzlement; 3) ambiguities and absurdities; 4) obscure means; and 5) anecdotes or personal experiences.

Explorations through reluctance to impose hierarchical relations between objects or events, reflected in their juxtaposition

Children's notions of objects and events originate from nature and imagination. Children are not overly familiar with conventional hierarchical relations between objects or events in traditional models, and tend to pose them in juxtaposition. In *The Philosophy of Childhood*, Matthews (1995) demonstrates a generally presupposed model for judging children's cognitive achievements (by adults). He presents an example of children's thinking patterns on *weight*. A model containing clear-cut measurements has to determine the right notion of *weight* objectively and reliably. A measurement in this respect could be a pair of scales to be read by an outsider. However, many children consider the weighing instrument and measured objects as similar things and suppose mutual relationships between the two. In their eyes, all objects are animated. An object like a set of scales may offer the results of the relationship experienced by itself; while at the same time representing results according to a presupposed model. So, on the one hand, the pair of scales is just a means to measure objectively; on the other they could be an entity endowed with feelings. It is posed in juxtaposition to the object being weighed and represents an independent outcome without any presupposed model (Matthews, 1994).

Explorations through questioning, wonderment, and puzzlement

Children ask many questions; they wonder about unknown things and puzzle about that which does not fit in familiar frameworks. Children ask questions to fill in 'gaps' in their knowledge exposed through their wonderment, and to exploit their puzzlement in their exploration of reality. The question "Mama, does a mess exist when nobody sees it?" (Chapter 1, introduction) is a good example. The girl simply explored the situation of her reality in the light of the presupposition of the initial demand to clean up the mess in her room. Exploring reality is questioning reality, is searching for the uncertainty of knowledge and the incongruities of experience. Wonderment and puzzlement are based on openness and function as doors into an infinite variety of possible interpretations, while at the same time making experienced events fit thinking patterns and *vice versa*.

The same effect seems to be realised by the use of special words or statements like *weird*, *odd*, and *strange*. Through using *oddness* or similar concepts, children display fundamentally open attitudes of wondering and questioning. Their judgement of 'odd' and questions like 'how is it possible?' do not seem to be intended to exclude objects or events of reference from a standard opinion, just to identify them as 'not fitting yet' or as incompatible at first glance. Weirdness is assigned to a particular object or event that may be made to fit into a reality framework and knowledge system. In contrast, adults usually refer 'odd' to an event that doesn't make sense or doesn't fit in their generally accepted concept of reality. 'Oddness' is often interpreted by children as being *something else*. For many children, quests seem to be driven by some magical manifestation of reality that cannot be explained and so they continue their exploration through trials, imagination and autonomous production of thoughts as instruments, using non-familiar meanings and strategies.

Exploration through ambiguities and 'absurdities'

When faced with questions that cannot generate definite answers, many children feel challenged to explore reality through unconventional strategies. Consequently they exploit ambiguities and absurdities to identify philosophical notions at the roots of reality, to clear arising contexts, and to sense alternative explanations. The correspondence between thoughts and reality and the supposed correctness of knowledge systems are challenged. Through exploration and these venturing procedures, not only the object of exploration is challenged but also the social environment with which it cooperates. Philosophical exploration becomes a joint venture with play-like features and arbitrarily formulated rules.

Philosophical notions are explored tentatively, in the search for ambiguity, vagueness, uncertainty and absurdities. Some concepts elaborated by children in the described discussions, appear quite absurd at first glance to adults. For example, the suggestion of a machine that can defy gravity; in the discussion about *Can the countries on earth change?* "… then you could, like, remove a country and put it onto Morocco;" and in the construction of a formula to make the bigger as if

observed through a microscope. A vague distinction between physical and the meta-physical world opens interpretations and explanations to absurdities. Contexts are tested by consciously presenting different interpretations. Conventional explanations are not taken seriously, but deliberately interpreted absurdly or in discordance with general accepted standards. Denying, reversing, or rearranging rules and generally accepted successions of events or thoughts is like a game with sense and nonsense. In her study about nonsense, Susan Stewart (1979) distinguished five strategies of nonsense exploitation, to make sense of nonsense: 1) reversals and inversions; 2) play with boundaries; 3) play with infinity; 4) uses of simultaneity; and 5) whimsical (re)-arrangement within a closed field.

These strategies may also show up in demonstrated thinking patterns on philosophical topics. Experiments often fail because children will turn rules upside down. Boundaries are blurred: for example, when asking to be half a vegetarian. Infinity is challenged by serial formation, as in the recurrence of *why?* or in the unlimited imitations of behaviour of one another (a very well-known behaviour among children). Incompatible meanings are exploited simultaneously and several contexts of one event are rearranged in the discussions highlighted in this study.

Many arguments explaining experiences or hypothetical events show absurdity being constructed from obscure, mysterious, non-standardised relationships, or show a *reductio ad absurdum*. For example, in a series of television programs on philosophising with children (Rondhuis-IKON, 1994), Niels, an 11-year old boy, was shown a picture of himself as a two-year old sitting with his father in a bath under the shower. A group of ten children, classmates of the same age, argued about what is real and what may be not. First, Niels tried to prove that this picture actually represented him: "It is really me! My mum told me. I loved water and ..." A classmate then interrupted him: "It also depends who and what you are in the picture with. If you are on your mum's lap, then you just know. But on your own, then you might not be sure." Niels continued: "No, but I don't remember that shower and that bath. I do remember that I always loved water and that I always went under the shower with my dad. I have got proof that I am sitting there with my dad, because my dad would not go under the shower with another child."

Exploration through obscure means

Exploring reality may go through conventional and intelligible strategies or methods; sometimes however, the strategies seem incomprehensible. Regardless, intelligible and incomprehensible paths may arrive at similar outcomes. Just as there is a sharp distinction between the justification and the good of decisions, so must *methods to arrive at outcomes* be distinguished sharply from their *outcomes* or solutions. Methods may encompass many types of reasoning. Some of them will be generally accepted, some of them will not. With respect to arriving at qualified outcomes, it seems difficult to distinguish between various reasoning methods. Mathematical thinking patterns dealing with numbers, serve as an excellent example. They are considered to be a realisation of clear and systematically constructed collections of figures in response to well defined problems with

exact solutions. Conventional procedures and solutions to mathematical problems and calculations on paper look identical for sets of problems. However, outcomes should be distinguished from the methods followed. Many children demonstrate non-conventional, peculiar, or idiosyncratic thinking constructions. They often select original entrees to complex questions, developing creative instruments unexpected by adults. Children's creative thinking in mathematics is perhaps underestimated. Some teaching methodologists in mathematics (e.g., Selter & Spiegel, 1997) and biology (e.g., Boerwinkel, 2003) have noticed remarkable reasoning styles among children and young adolescents. Some might not be well understood by their teachers because the reasoning styles do not fit into a particular, generally accepted model. This may be demonstrated by the method Sven (age 7) followed when he was asked to add the following numbers: 9, 12, 10, 11, 8, 10, 9, 8, 12, 11, 10 and 12 (Selter & Spiegel, 1997). Sven displayed the following thinking pattern: 119, 121, 121, 122, 120, 120, 119, 117, 119, 120, 120, 122, and asked the teacher if this was right. The outcome of 122 is right. Sven's thinking strategy however, consisting of the foregoing numbers, is not immediately clear. First, he made an estimation of the average number of numbers to be added; subsequently he traced out the difference between each number and this average number, and finally arrived at his outcome.

Explorations through narrative events

Reality arises as a series of situations and events that do not seem structured conventionally in advance. Exploring this unfolding sequence elevates experiences toward thinking patterns; compares them with other experiences; and enables a child to communicate them. Anecdotes seem to be exploited as touchstone of conceptual understanding. The procedure of deriving conceptual relationships from concrete experiences is called 'regressive abstraction' (Nelson, 1970). With children, it goes often through unconventional analogies, associative lines, and intuitions that are not immediately intelligible. Oppositely, a smooth transition from abstract thinking levels toward concrete occurrences can be explored through narrative events: personal experience as realisation of conceptual understanding. In response to problematic questions, children present anecdotes on several everyday experiences; instances are sensed and compared with each other, leading to the discovery of content and boundaries of concepts.

Demonstrable thinking patterns, driven by philosophical quests seem to force many children to explore reality, ending with a display of autonomously generated thoughts, including many unconventional turns and nonsensical modes to stretch existing norms or requirements of correctness. The tendency to make sense of nonsense is adventurous and stimulates creative thinking. Many reasoning styles displayed by children do not fit generally accepted systems of reasoning and problem solving. They may derive from incoherent world-views, from absurdity, or from wild associations in response to experience. However, these thinking patterns can be detected as working strategies to make experience fit theoretical frameworks, to reach clarity and qualify as being philosophical. To identify and detect philosophically qualified thinking patterns seems to only require the

appearance of but a few characteristic expression styles. The reluctance to impose hierarchical relations, questioning, wonderment and puzzlement, ambiguities and absurdities, obscure thinking methods, and narrative events come to the surface without transparent relationships to the explored themes.

In this study, children's thinking patterns incited by philosophical questions will be collected and conceptualised as a form of intellectual disquietude through expression styles that reflect numerous intangible notions detached from reality. The presence of such expressions will be assessed without the application of conventional criteria used to evaluate philosophical discussions.

4 Measuring thinking patterns

In this chapter, the emphasis is on typical thinking patterns shown in philosophical discussions amongst 11 to 12-year olds. Children in this age group show qualities that reflect the main features of philosophising gathered from philosophical literature (Chapter 2). The philosophical quality of thinking patterns might be considered as elusive or as a essentially non-measurable theoretical construct. However, it is the aim of this project to capture this divergent character in a few indices. In Section 3.2, it was demonstrated that conventional criteria and evaluative approaches are inadequate when it comes to identifying philosophical qualities; nevertheless, the discussion samples represent philosophical thinking patterns that have found concrete shape in behavioural expressions. Many of these expressions can be detected and empirically investigated. To this end, thinking patterns must be captured in standardised situations first, generating a database that can be subsequently processed quantitatively. Philosophical characterisations and psychometric methods will complement each other instead of being poised in opposition.

Philosophy and psychology

As a purely theoretical study, philosophy focuses on themes and their development. Philosophers deal with concepts. It is the concept of a horse, the concept of time, life, or knowledge, embedded in their frameworks that is uncovered and researched. The guiding principle is the inquiry by itself. To provide an estimate of the philosophical qualities of children's thinking patterns involves beginning by looking at their thoughts, searching for presuppositions, constructing solid lines of thoughts, inventing alternative thoughts, and comparing all imaginable thoughts, irrespective their actualised or fictive nature, and irrespective their carriers. Although theoretical approaches suppose encompassing real-life experiences, the discipline of philosophy usually avoids *measuring* philosophical characteristics, for fear of falling into contradictions. Philosophical reasoning endeavours to 'think the whole' venturing beyond experience in order to answer questions about the origin of knowledge and thinking patterns. So it seems impossible to measure unanimously or quantify experienced philosophically qualified thinking patterns by philosophical means.

Psychology is an empirical science that focuses on individuals and their development. Psychology deals with observable behaviour and its carriers, thinking, perceiving and feeling human beings. Observed relations between behaviours are mapped onto models reflecting a theoretical structure of supposed properties. Sometimes, such a model is as simple as a scale. Thinking patterns will be approached as observable behaviour. In psychology, evaluative approaches are very common, using criteria formulated *a priori* or *a postiori*. Some approaches within psychology tried to avoid evaluative criteria. Piaget, for example, tried to avoid *a priori* criteria in observing behaviour and framing his stages of development. So did Kohlberg (1981) in his studies on moral development. On a fundamental level, these studies do not reckon with the status of 'thinking patterns', which ground tendencies of enquiry into the unknown, and which are sensitive to ambiguity, vagueness and uncertainty, and allow the autonomous production of fresh thoughts.

The status of philosophically qualified thinking patterns

What about the status of philosophically qualified thinking patterns? Thinking patterns concern an activity or property of the mind, referring to thoughts as their sediment. Accordingly, thinking patterns should be recognised similarly or as meta-constructions. They will be philosophically significant if the main philosophical tendencies are met: sensitivity to ambiguity, vagueness and uncertainty, reasoning qualities, systematic inquiry into the unknown, and construction of relations between different domains of knowledge and experience. Thereby, these tendencies are concerned with the origin and acquisition of knowledge, and refer to epistemological questions about the possibility of knowing and thinking. These questions have inspired many to attempt to represent knowledge, including philosophers like Kant and empirical researchers like Piaget. Both scientists were faced with the same observable events (production of knowledge) and were equipped with the same culturally derived capacity to create theories. However, their focus differed. As a philosopher, Kant reasoned logically about knowledge into a Copernican turn (Copernican revolution in philosophy). His solution was to trace the boundaries of the proper uses of reason. For him, this meant the construction of twelve categories of pure reason delimiting the boundaries of possible experience. The psychologist Piaget constructed a theoretical framework for knowledge acquisition in experimental terms. His solution was to trace the boundaries of knowledge according to observed epistemological developments (Chapman, 1999).

Measuring philosophical quality

This interdisciplinary study examines observable expressions of philosophically qualified thinking patterns. The main features of philosopher's activity are approached by psychometric methods. However, potential philosophical quality is not measured using conventional criteria concerning maturity or correctness, or through the judgement of the legitimacy of thoughts and concepts. Thinking patterns expressed by children will be judged on the presence of observable and

countable events, identifying the elaboration of philosophical themes and notions by children, according to a theoretical construct of philosophising. These events will be observed, collected and counted with regard to the number of expressed thinking patterns such as: questioning, hesitations, stretching conventional meanings, using anecdotes, special words, new expressions, metaphors, and so on. Qualified expressions derived from philosophical discussions, like those demonstrated in this chapter, will be identified and collected. This collection of expressions will constitute a database that subsequently will be processed quantitatively. All of this is reported in Chapter 4. Quantitative and qualitative analyses may detect a philosophical quality, referring to a supposed individual cognitive competence or ability. Although some of the main philosophical features might be captured using the proposed approaches for measurement, the procedure might not cover all of the philosophical expertise.

5 Conclusion

To bridge the gap between the philosophical theoretical approach of Chapter 2 and the empirical approach of Chapter 4, six fragments of actual discussions with 11 to 12-year old children were examined. These discussions, generated in response to a variety of philosophical questions, were contemplated and commented upon. Children's explorations of philosophical questions involve tentative behaviour or trials to find answers. They are driven by an intrinsic force to 'capture' the unknown and will proceed along their own non-predictable and whimsical way. This characteristic makes them difficult to judge using conventional criteria. Philosophical qualified thinking patterns seem to elude evaluative norms like presuppositions of maturity, requirements of correctness and cultural justified agreements. Vagueness and ambiguity are generally accepted as the pre-eminent drive to evolve philosophically qualified concepts and ideas. For this reason, the expression of thinking patterns will be examined as openly as possible, yet still remain grounded in the three main tendencies of philosophy.

The bridge constructed between philosophical and psychological approaches consists of eliciting thinking patterns as observable behaviour in standardised discussion situations. Discussions must be standardised in order to make children's performances mutually and temporally comparable. These performances are treated as observable occurrences of materialised philosophical main features. This materialisation takes the form of specific indicators, each scorable for presence/absence, to be presented in Chapter 4. So, observable events in the apparently boundless elaboration and infinite paths through themes will be formatted. Through these formats, philosophically qualified thinking patterns can be measured empirically and processed quantitatively.

4 *Assessing Philosophical Quality*

Observations of youngsters exchanging their philosophical thoughts drive researchers to search for thinking patterns. A scientific approach requires a conceptual framework to elicit the philosophical quality of expressed thoughts preferably based on the main pillars of philosophy and wisdom. In this study, six indicators are proposed to assess the philosophical qualities of utterances by youngsters (Section 4.1). A 'tetralogue' is developed as a standardised situation in which philosophical questions are discussed with four youngsters (Section 4.2). Their video- and audio taped utterances are transcribed, formatted and checked on presence of indicators of philosophical quality (Section 4.3). In the results (Section 4.4) the objectivity (Section 4.4.2) and reliability (Section 4.4.3) of this procedure will be investigated firstly. Next, two numerical indices will be constructed: one that reflects the philosophical quality of an individual (pq) performance (Section 4.4.4), and one that reflects the philosophical quality (PQ) of the tetralogue as a group performance (Section 4.4.5). Finally, the outcomes of pq indices and PQ indices will be compared with other findings on individuals and tetralogues to test the validity of pq and PQ (Section 4.4.6).

1 PQ Indicators: a conceptual framework

In Chapter 2, three main pillars of philosophical thinking and wisdom were revealed: one stressing analytical and reasoning qualities; one concerning the dealing with ambiguities, vagueness or borderline explorations, and uncertainty; and one focusing on the contact with real life experiences.

In this chapter, the focus will be directed on the observable carriers of philosophical qualities of thinking patterns: utterances made by individuals and by groups of four youngsters in standardised philosophical discussions called 'tetralogues' (tetra = four, gr.).

An utterance is defined as the uninterrupted train of *expressed thoughts*, or the continuous flow of sayings by one of the tetralogue participants. First, it will be argued that these utterances reflect mental activities. Each utterance consists of many smaller parts of actual expressions that are scored separately. Secondly, a situation must be created as a precondition for the autonomous production of a systematic train of thoughts in such a way that identifications of philosophical qualities are comparable between situations and individuals. Also, a conceptual framework must cover the theory of philosophical tendencies and real life expressions in tetralogues allowing the assessment of philosophical qualities of thinking patterns. The merger of theoretical analysis and the actual expressions of thoughts in utterances results in six indicators as the manifestations of philosophical thinking patterns. These indicators, adequately combined, serve as promising tools to assess philosophical qualities of youngsters and of tetralogues.

Utterances and thoughts

Many researchers, even behaviourists like Skinner (Verbal behavior, 1957), agree on the fact that utterances do reflect mental activities. Others, philosophers like Quine (1972), are departing from this concept. Since mental states and processes cannot be observed directly one cannot verify the correctness of translation of private experienced mental activities into the linguistic performance of these activities; thoughts and thinking patterns are not necessarily reflected in words. Meta-linguistic awareness or the ability to think about language demonstrates the distinction between thoughts and their translation in words logically, the distinction between linguistic competence and linguistic performance (Chomsky, 1957), the distinction between object of reference, representation, and meaning (Peirce, 1992). An argument in favour of the presupposition for the coincidence of thoughts and the linguistic performance of them is given by cognitive developmental inquiries (Berk, 1997). Developmental psychologists consider comprehension as recognition of meaning, and production as recall of words and concepts (Berk, 1997). All cognitive activities, like comprehension and production of concepts, the receptivity to language, the expressive style of mastering language, the under- and over-extension of word use presuppose a direct link between thoughts and their translation in words. Successful communication between people may demonstrate the justification of this presupposition. Linguistic utterances are used as identifying marks of mental states and processes (D'Andrade, 1989). Therefore, it seems reasonable to assume some correspondence between thinking patterns and expressions of thoughts. This may give way to identify potential indicators for determining thinking patterns from expressed utterances.

Manifestation of philosophising

Utterances with relevant modes of expression, revealing philosophical tendencies, can appear in situations that are directed by a philosophical quest in a community of inquiry (Lipman, 1977). Secondly, expressions and utterances must be accessible for comparison: individual contributions mutually, and tetralogues with other tetralogues (i.e., the situations must be standardised). This standardisation will be demonstrated in Section 4.2. Finally, expressions and utterances will be examined and analysed through a limited set of features based on the main pillars of philosophy. For this purpose, a pilot study of six philosophical discussions with youngsters was designed. Each discussion contained about 250 utterances, each containing a variety of expressions. Expressions and meanings were examined within the context of the set of utterances in which they appear; evaluated as to whether they were in concordance with the main pillars of philosophy and findings of wisdom investigation; and then compared mutually (Rondhuis & van der Leeuw, 2000). Close observation of these utterances is shown to reveal various modes in which youngsters try to express thinking patterns.

How do youngsters exploit ambiguities, uncertainties, vagueness and borderline explorations? How do they deal with the sensation of real-life experiences? In what terms are they performing regressive abstraction, reasoning and analysing qualities?

What do their trains of thought look like? One cannot simply decide to determine uncertainty by only a sneaking suspicion. Moreover, a philosophical discussion with children or youngsters seldom has a clear-cut direction, the participants do not always finish their lines of thought, and interaction has a big influence on the course they follow and their way of expression. On developmental grounds one may quote that the unavoidably rapid changes children and adolescents experience during the process of maturation and simultaneously the underestimation of their knowledge, belief and judgement, measured with adult standards, undoubtedly generate many of so-called imperfections in their linguistic utterances or expressions that at least are difficult to judge. Finally, if just specific linguistic utterances are scored, we have to be sure they are intended, and not simply used as fillers or automatic reactions.

Can an ape be an artist?

Nataly: Yes, but if an ape, well ... Suppose ... because, he just said that an ape is an artist at the very moment that he finished the painting, at least ... you just said, didn't you say so? (Nataly is looking at Michiel) Michiel: Yes. Yes, well, but ... suppose that ... a tiger is doing it. Is it (even) art by then? Or if a ... snail is doing it. Is it art, just then? Michiel: If a tiger makes something that people like, does this mean that the tiger is performing art? Nataly: Yes, but thus ... so ... Kerin: It depends also ... So, if people think it is beautiful, is it consequently art? Michiel: Yes, if people like ... if ... if men and women think it is beautiful. Arthur: It is always what it was with arts ... But ... Kerin: Yes. But what is then ... Arthur: What is art anyway? Kerin: Yes ... Arthur: So you are not ... Nataly: Yes, but there are many pieces of art that I personally do not like at all. But they are still pieces of art. Michiel: Yes, personally. But in general people who think they are beautiful are people who do have money to spend for them. And consequently they call it art. It doesn't matter what you think, those things are still called pieces of art. Arthur: Some time ago, they had an elephant ... they have been started in Thailand with elephants, to give them a brush just in their proboscis, and ... Well they are drawing and painting as well, five or six elephants ... Kerin: Well, I mean, that is not voluntary. I think that ... yes, I don't know, it may be stupid: but only art on a voluntary base ... But I do have really something like ... yes, those animals are used very much ... Michiel: If I would be exploited ... Arthur: They can simply choose their own colors. They can simply ... Michiel: If it is not voluntary, then it becomes ... Kerin: No, that is not the case ... Michiel: ...that does not refer to the question whether it is art. Keren: You are right, but you are stressing that he is an artist in those cases. It appears to me that you have to choose to be an artist by yourself and that you are so by yourself. Nataly: Yes, but being an artist, does it depend of being a human or of the voluntariness of the activity? I mean ... for it is just still possible

that ... some animals are doing it voluntary. Are those animals artists by then? Michiel: Mmmm Kerin: Yes. That is a good question. I made paintings just as well. Michiel: Yes all right, but those were not fantastic. Nataly: Yes but when, when (laughing) when is something a piece of art? I mean, this is also something to carry on. Is it a piece of art if everybody likes it or is it a piece of art if ... if ... Arthur: If it is different ... Nathaly: ... if it is different? Or is it a piece of art if it carries a certain name? Or is it a piece of art if it contains certain colours ... Arthur: Or if a certain person ... Or if it is representing a certain picture, or something ... Yes, when is something a piece of art anyway?

Fragment from a tetralogue with 4 participants of 16 years old, two boys and two girls, about the question: can an ape be an artist?

Serious difficulties in inferring specific thinking patterns from observable expressions are encountered through two opposite moves: one departing from youngster's utterances, the other departing from the theoretical founded philosophical features. These moves are expressed on three detectable levels of allocation of philosophical thinking patterns: concrete materialised (para-) linguistic expressions, their meaning with allocated meaning components, and indicators of philosophical quality (see Appendix II and fragment of it, Table 4.a). Moving from practice into theory, beginning is realised by utterances, containing a variety of actual expressions. Concrete materialised linguistic expressions used by the youngsters, reflecting similar philosophical meanings were put into the same row and collected as philosophical quality indicators linked the main philosophical pillars. Moving from theory into practice, initially, philosophical pillars were translated into a variety of categories representing philosophical tendencies in different ways. Through grouping of meaning components, these tendencies were expanded literally with specific verbs, others were materialised in combination of words and attitudes with clear meanings, turns of phrases, dictions, moods, etc. Following 'Grounded Theory' (Glaser & Straus, 1967), once indicator categories were saturated with all possible meaning groups and no new phenomena of expressions appeared, six philosophical quality indicators were identified and assessed.

The modes of expression in which a philosophical quality could be detected were identified and depicted on a variety of levels of language: on the level of words, expression modes, syntactical moves, non-verbal although meaningful sounds, or combinations of these. The simplest form is the use of specific words like causal conjunctions or cognitive verbs. The use of the word 'because' is indicating a reasoning quality in most cases. There are also combinations of words and modes of expressions. For example, 'I think' shows an epistemic position. The negation of an expression of knowing often is referring to an uncertainty. The use of metaphors or other modes of comparing may indicate a reasoning quality and or an expression of tentative behaviour. Open questions apparently show openness. There are unusual syntactical moves demonstrating a thinking quality of being 'out of order' or the use of modal verbs or connectives, for instance containing

Table 4.a. Part of the register translated from Dutch (Appendix II): characteristics of indicators on three levels.

Level 1: Linguistic expressions	**Level 2: Meaning components**	**Level 3: Indicators**
To look like	Ambiguity	Idt
To seem, to appear	Being uneasy about	Idt
	Linked to 1 singular or plural	Ep
	Description of a resemblance	-
To appear	Reasoning function	Re
	Assessing (neutral descriptive)	-
As it were	Ambiguity	Idt
	Tentative trial	Te
	Stopgap, filler	-
Seeming(-ly)	Ambiguity	Idt
Apparent(-ly)	Relativising, putting something in perspective, sepsis	Idt & Ep
Obvious(-ly)	Tentative trial	Te
	Reasoning function	Re
To call, to be called,	Ambiguity (excepting: 1 singular)	Idt
To consider (regard) as,	1 singular (I call…)	Ep
They say…	Assessing	-
If you look/interpret from a certain angle …	Relativising, putting something in perspective, sepsis	Idt
'between quotation marks'	Quoting	-
	Relativising, putting something in perspective, sepsis	Idt
As if	Vagueness, ambiguity	Idt
	Comparison in function of reasoning	Re
Real(-ly), actual(-ly),	Ambiguity, borderline exploration	Idt
fundamental(-ly),	Discriminating the non real	Idt
Honestly spoken	Tentative trial	Te
In fact, in principle	Stopgap, filler, expression of prudence	-
At first glance	Linked to a cognitive verb 1 singular	
	The real thing, true	-
	Emphasising or stressing a meaning	-
	Stammering, stumbling, postponing	Te
You don't have a real proof	Uncertainty, relativising	Idt
(Not) Everyone does see this	Uncertainty, relativising	Idt
A little, half	Ambiguity	Idt
Vague	Discriminating	Idt
A kind of, a way of	Tentative trial	Te
More or less	Expression of prudence	Te
Something like that, some …	Stopgap, filler	-
It 's almost improvable		
Difficult	Not knowing, detecting an ambiguity	Idt
	Wondering	Op

Idt: Indecisive thinking; Op: Openness; Te: Tentative behaviour; Ep: Epistemic position; Re: Reasoning quality; An: Anecdotal quality.

a possibility or modality. Verbal attitudes like recurrences, hesitations, incomplete sentences often indicate a tentative behaviour. It may also express a combination of uncertainty and willingness to discover. Verbal expressions of mood or sentence types, like an ironic way of expressing, may be very important. Intonation, extra-linguistic gestures, and contextual features may also betray a specific thinking move. Of course, a direct reference to explicitly vague and ambiguous aspects of events or to life experiences is an obvious indication of receptivity to ambiguity. Often these tendencies are concealed in a variety of semantic, pragmatic, and syntactical features. So, indicators to identify thinking patterns, appearances of wonderment about the world and speculations to make sense of it, differ in quality, are operating on various linguistic levels and are not detectable in one dimension only.

Appendix II contains a list of such expressions in Dutch with their supposed exploited meanings and how they were assigned to philosophical quality indicators. A preliminary edition of the appendix was formulated after evaluating six pilot discussions and constitutes the base of 'Performance and progress in philosophy' (Rondhuis & van der Leeuw, 2000). The grouping process was began with the utterances from the six pilot discussions and was continued with analyses of utterances from the main study until saturation of the meaning groups occurred, i.e. no new expressions were found. In its final form, Appendix II comprises observations, registering, and evaluation of some thirty philosophical discussions.

Considering all previously mentioned aspects, six indicators for philosophical quality were distinguished, correlating with six thinking patterns or attitudes representing specific philosophical qualities or aspects of them. Each indicator covers a group of linguistic expressions. These indicators are:

1. Indecisive thinking (Idt)
2. Openness (Op)
3. Tentative behaviour (Te)
4. Epistemic position (Ep)
5. Reasoning quality (Re), and
6. Anecdotal quality (An).

Indicator 1: Indecisive thinking. This philosophical quality indicator points to the discovery of ambiguities, contrasts and uncertainties. It is the most complicated indicator because it shows itself in various ways. Sometimes it is demonstrated by means of modal expressions, or in expressions of disjunction, of separation, or of degree. Sometimes borderline cases are explicitly made questionable. Often children use expressions like 'fundamentally', 'in principle', 'actually', 'at first glance' or verbs like 'to seem' and 'to call', indicating a distinction between appearance or name and reality. The indicator of *indecisive thinking* refers to 'the other side of the picture', to the unknown, showing the awareness that events can be understood in various ways. Sometimes youngsters refer to the fact that 'it is difficult …', frowning their brows, or they refer to contra intuitive ideas. In many taped discussions, ten- to twelve-year-old children regularly use expressions like (in Dutch) 'eigenlijk', meaning something like 'really', 'in fact', 'honestly spoken' or 'well,

ordinarily I would, but ...' (Rondhuis, 1994, 2001). According to Morris *et al.*, (1994) the latter expression is used as a modal expression indicating that one is generating doubts and has a feeling for the ambiguities and uncertainties in what was said previously. At the same time, the expression is no definitive judgment, but leaves open several possibilities. Frequently, children think in terms of 'that's crazy'. The use of words like 'crazy', 'odd', 'strange', 'peculiar', or 'difficult' is a real demonstration of uneasiness and an emerging awareness of the multiple ways to understand events. There is so much youngsters are supposed not to know yet. The use of the word 'crazy' makes one wonder about things that are not known or events that are not yet fitting in their present framework. The expression signals an urge to look for a wider framework in which events do fit.

Indecisiveness and uncertainty can also be expressed by a variety of words of not knowing and substitutes for not being sure: 'somewhat', 'more or less', 'it depends', and ways to express that potentially there may be alternatives. Although the name suggests otherwise, patterns of indecisive thinking do not point to the inability of making decisions but to the discovery of vagueness and relativity. Demarcations between different concepts for example appear to be made arbitrary and more complex or paradoxical than it looks like at first glance. This indicator characterises an intellectual inquietude.

In the pilot approach and on theoretical grounds a contra-indicator of indecisive thinking was formulated (Rondhuis & van der Leeuw, 2000). This indicator betrays the existence of definite views, certainties, and an appeal on authorities and was to be linked to Idt as the non-appearance of indecisive thinking patterns. Authorities can be of various types: powers, rules, majorities, standard opinions, or sometimes evidence from the senses. Certainties are definite judgements, which claim universal validity. The contra-indicator was not incorporated in the final list of indicators of this project, since the philosophical discussions were standardised and all introduced in the same way, each comprehending a simple and comprehensible definition of what philosophising may be and a simple and comprehendible set of rules to perform philosophising. One of the rules of the tetralogue is the performance of thinking by yourself, and consequently the radical prohibition of authority's arguments, based on authorities from outside, i.e. apart from the senses. Nevertheless, when these utterances appeared, participants corrected themselves, or were attacked and rectified by others. Sometimes the energetic use of expressions like 'it is required, forbidden, and impossible that ...' may look like contra-indicators. But in most of these cases the performance is a demonstration of a need for certainty often followed by an argument.

Indicator 2: Openness. Indicating a philosophical quality, openness is not a personality trait but a state of mental activity. This indicator is demonstrated in expressions of wonder, in the apparent readiness to meet the unknown and break open one's view without being forced by clear economic benefits, physical inevitability, or psychological pressure. Openness implies the disinterested acknowledgement that some thought had not occurred earlier to someone, and secondly, attempts to corroborate a specific line of thought. Finally, some kinds of recalcitrance can be identified as special types of openness in the sense of

refusing to accept any view without further ado. In strong denials or negations, this stance may be recognised as the intention to ferment trouble, in challenging ways of expressing, and in provoking contradictions. Some utterances, however, apparently show openness but prove to be an instrument or a strategic move with a determined purpose: for example the seemingly open question 'what else is there to believe?' pronounced with an intonation suggesting there is no other possibility for believing. If this expression is used as an argument to emphasize one's opinion, it should be understood in a train of reasoning rather than as a token of openness. This example makes clear that merely examining word meanings or the meanings of descriptive expressions is not satisfactory. Questions for clear-cut information do not count as examples of this indicator. These questions are tight and closed. They are not indicating openness. Openness is a divergent thinking pattern. The indicators of indecisive thinking and of openness play a role of the same kind in the entire design of evaluation and calculation of philosophical quality.

Indicator 3: Tentative behaviour. This philosophical quality indicator points to an attitude of trying-out. Contrary to openness, tentative behaviour denotes mainly a convergent thinking pattern. There are three types of tentative behaviour: a conceptual, a semantic, and a more formal type. The conceptual type comprehends the conceptual trials like posing a contra-factual hypothesis. The semantic type may be recognised in linguistic attempts like the use of unknown or no conventional words, an original flash of intuition, unusual comparisons, sometimes by means of metaphors, or strange combinations of phrases like in 'it has a much bigger impossibility'. Semantic tentative behaviour always implies a search in words for boundaries, contrasts, and new frameworks of meanings. The more formal type may be recognised in stammering or stumbling, imperfections, unfinished declarations, irregularities and anomalies, abrupt (associative) switchovers, often with an idiosyncratic quality. Since incorrect and strange word uses are not clearly distinguishable, formal and semantic tentative behaviours are fluent. This may be shown in some unusual comparative expressions, pleonasms, contaminations, non-intentioned oppositions and so on. Tentative behaviour is also displayed by illustrations, and the use of examples. Finally, tentative behaviour is demonstrated in expressions of prudence, in the use of special terms such as 'as it were', 'something like that', sometimes attention holding devices like 'oh, now ...', 'well, look ...' or in the conscious use of nonsense and unconventional ways of reasoning. Both the indicators of openness and tentative behaviour are responsible for the inquiry character of a philosophical dialogue.

Indicator 4: Epistemic position. The fourth philosophical quality indicator points to the epistemic position a speaker tends to hold with respect to the propositional content of his utterance. The epistemic position represents an attitude toward knowledge and experience. It is a demonstration of certain detachment between speaker and his declaration, normally effected by means of performative formula (reflective and cognitive verbs used in the first person singular, present indicative, active). Typical expressions for this position are: 'I think/believe/wish/... that', 'according to me', and utterances of not knowing, doubting, reflecting, deliberating, realising, contemplating, or imaging. Some ironic or cynical phrases may also betray this position. A demonstrated detachment

between speaker and object of his speech may depict an intellectual distance characterising a philosophical quality. In those cases, performances of thinking and the like are followed by performances of indecisive thinking, tentative behaviour, or demonstrations of reasoning. When an epistemic position appears alone, the expression seldom is an indication of the so-called intellectual distance. It is doubtful whether the use of cognitive verbs in the first person singular is intended to point to the relation of the speaker to his experience of reality. Often the 'I think …' is meant as an expression of prudence more often used by girls than by boys (see Chapter 6). Role taking, substitute or vicarious thinking, feeling or valuing are also demonstrations of an epistemic position. Nowadays it seems to be fashionable to express yourself starting with 'I think …', referring to moral aspects of the quest. Even in philosophical quests about the demarcation of observable sizes, such as small and big, pupils became the victims of moralistic restrains and refer for instance to the moral equality between people irrespective of their size.

Indicator 5: Reasoning quality. The relevance of reasoning in thinking moves stands out clearly as an indicator for philosophical quality. Analytical and reasoning qualities - no matter whether argumentation is correct or false - may be recognized in the use of sentences with a specific structure (e.g., 'if-then' sentences) or of specific words, like 'because', 'for', 'if', 'thus', etc., in logical thought-schemes, in making reversals, or in analogical reasoning. The use of an 'if-then' sentence comprehending a contra-factual hypothesis followed by a consequence is indicating tentative behaviour and a reasoning quality as well. When the 'if '-part of the 'if-then' sentence realises the condition of the consequence, only a reasoning quality will be assigned to the expression. The combination of tentative and reasoning behaviour may not only be found in 'if-then' sentences but also in many trials of comparing and upholstering. As mentioned earlier, it often proved difficult to make a strict distinction between different indicators. For example thinking in analogies and thinking by upholstering are theoretically clearly separated, but both are thinking patterns based on the mental operation of comparing. In many discussions with youngsters, it is not possible to distinguish whether the speaker is evoking images of a general comparative nature, or referring to real-life pragmatics. In both cases, they try to trace boundaries by discovering differences and similarities. Thinking in analogies will be combined with an epistemic position, in an indirect way of speaking, while referring to real-life instances may occur in a very direct, maybe less conscious way. Sometimes the reasoning quality – also called informal ways of reasoning, is expressed in rhetoric's like in a rhetorical question or in a regular question followed by an indicative sentence. A very unusual, maybe paradoxical performance of reasoning is the expression 'so it must …', when a speaker concludes to a judgement.

Indicator 6: Anecdotal quality. This philosophical quality indicator includes upholstering concepts or ideas with real-life experiences, in addition to finding connections between abstract ideas and concrete situations. Recognition of these connections shows up in declarations about experiences the speaker has had, and references to specific places, times, objects, and persons. The value of anecdotes in philosophical discussions is disputed. Analysis of daily experiences, according to Nelson (1970), or anecdotes is the basis of discovering the conceptual framework

of the speaker underlying his discourse presentation. Alternatively, some reports about philosophising with children consider a great portion of anecdotes unfit because discussions are likely to get bogged down in babble (Vandaele, 1996). Recounting anecdotes results often in the use of many words, especially by comparison with the very elliptic ways of expressions youngsters are used to. It also proves infectious. Primary school children use anecdotes more often than those in secondary schools. Sometimes anecdotes are mere references to one's own life or an escape into case histories; sometimes they are of a more general character. Anyhow, this indicator will depict a philosophical quality only if it illustrates mediation between the tangible world of a concrete story and the non-tangible world of ideas. Consequently, it will be valued in the context of the entire discussion, only. Single performance of an anecdotal quality is not of a specific philosophical value in its own right.

2 The tetralogue: a standardised procedure for determining *philosophical quality indicators*

2.1 Introduction

Following the theoretical descriptions above, in this section, *philosophical quality indicators* are identified in real-life discussions collected from youngsters to understand the philosophical quality of their thinking patterns. This requires an instrument or method to gather relevant utterances of youngsters. These data should be collected in a systematic way in order to make it possible to compare the performances of participants in discussions over time and in different situations. This process requires standardisation of the situation in which utterances are generated, the procedures, the registration, the transcription, and the assigning of *philosophical quality* indicators to the collected data to ensure that results are objective, reliable and valid (Drenth & Sijtsma, 1990).

Utterances are generated in a standardised discussion forum called a 'tetralogue'. The philosophical quality of youngsters' thinking patterns is ultimately derived from standardised, (video-) recorded and transcribed philosophising sessions, followed by the assignment of *philosophical quality* indicators to the transcriptions, and the data-processing (Section 4.2). Objectivity and reliability of this process are evaluated in Sections 4.4.2 and 4.4.3, respectively. In Section 4.4.4 and 4.4.5, two numerical indices for philosophical quality are constructed reflecting the philosophical quality of an individual's contribution to a tetralogue and that of the dialogue between tetralogue participants as a whole, respectively.

As to the philosophising sessions, the preconditions should be equal in each session. This applies to the composition of the group of participants and to the session procedures (chair, topic selection, duration). Based on 15 years of experience with philosophical work with children, and on the results of six pilot experimental sessions ahead of the present study, the size of the group in a session was fixed at four youngsters. This number proved to be convenient both in terms of generating

adequate interaction by exchanging arguments on the one hand, and maintaining sufficient mutual attention among the participants. Recruitment of the session members is discussed in detail in Chapter 6.

In a standardised session, the discussions among the four youngsters were led by an adult expert and are spurred by a philosophical quest. In this study such standardised sessions are called 'tetralogues' referring to interactive discussions between four people of comparable level ('tetra' is four in Greek and 'logue' as an extension of dialogue). Tetralogues fulfil Lipman's (1977) most relevant requirement, i.e., transforming the group of youngsters into a community of inquiry. For the present study, about 100 tetralogues were conducted, including a series of follow-ups to check subsequent *philosophical quality* development among a group of four schoolboys over two consecutive years.

2.2. Session & Proceedings

Preparation and logistics

Schools and organisations for guiding youngsters were found to be the best locations logistically and physically, for conducting tetralogues since they usually provided standard facilities in terms of accessibility, rooms familiar to the youngsters, and infrastructure like black/white boards, appropriate furniture, space, light and electrical power for video recording. Schools were selected according to the required population of youngsters in terms of age, educational levels and regularity of life course (see Chapter 6). Of the 21 institutions approached, only one refused, and that was because of an overbooked programme. Selected schools or other organisations were invited to cooperate in this research while outlining the aim of the study, the voluntary nature of participation, and session procedures.

The collection of data was executed in three steps. A plenary meeting was scheduled prior to the standardised session to inform, explain and prepare the entire school class population. The information concerned the aim of the inquiry and the presence of a video camera. Usually the author was questioned by the pupils for their chance to be on television. The clear 'no' answer appeared to be a disappointment in almost all cases. Furthermore, this meeting included explanation of the nature of philosophising, the tetralogue procedure and rules, and the type of philosophical questions to be discussed. In the plenary meeting youngsters were invited to consider which type of philosophical questions would interest them most, on the condition that all participants had equal accessibility to the questions, and that these could not be provided by a final answer but at the same time would allow a systematic search for answers.

Next, the class was split into groups of four youngsters, preferably with mixed compositions of girls and boys. The number of such groups depended on the number of volunteers. 0 to 10% of the pupils did not volunteer. In some cases these volunteers and groups were pre-selected by the school teacher. Subsequently,

each group determined its preferred and most challenging philosophical question, complying with the boundary conditions set before. The second step consisted of the standardised session, the tetralogue. In this session the preferred philosophical question was extensively discussed during approximately 40 minutes. The tetralogue was recorded on video and audiotapes.

The third and last step consisted of the participants completing evaluation, personality and 'non-verbal intelligence' questionnaires. Occasionally, the three meetings took place the same day. Normally, the three meetings covered a time span of several days but never exceeding two weeks. Sometimes enthusiastic pupils were allowed to participate in another session as well, avoiding the first and last step of the procedure.

Group composition

Boundary conditions for tetralogues included that youngsters participate entirely on a voluntary basis, that participants belong to the same age category and to the same educational level.

The chair

A chair-person is responsible for conducting the performance of the tetralogue. Out of 95, 75 tetralogues were guided by the author; 17 by Philosophy students; 2 by students in Psychology; and one by a regular school teacher. The students and teacher received instructions prior to the sessions. Interventions from the chair were restricted to three types:

a. Intervention 1, questioning for clarification in case of (possible) misunderstanding.
b. Intervention 2, reiterating previous remarks and questions.
c. Intervention 3, quoting a fresh illustration of the topic in the rare cases tetralogues would require such impulses.

As for reiteration, three types of situations were noticed in practice: when part of philosophical question was not discussed, when the discussion drifted too far from the selected topic, or to confront statements with earlier, and possibly contrasting ones. This last form of reiteration often proved rather constructive. Interventions offering a fresh anecdote to illustrate the question at stake only occurred in rare cases when a tetralogue threatened to fail. In general, the attitude of the chair should be open, wondering, motivating and preventing the opposition of 'no, yes', or debating in general, by emphasising that 'nobody knows the definite answer'.

Checklists of tools, instruments, consumables, times and activities were available to the chair for each session.

Philosophical questions

In all cases, the philosophical question acting as the starter of the tetralogue had been chosen and formulated by the youngsters themselves. This question was embedded in a vivid, detailed and concrete story or report. Concrete elements in the question were needed to prevent a knitting mess of words and thoughts that could run away wildly. Such requirements are also requirements for a Socratic debate (van der Leeuw, 2005).

Questions were of an open nature, excluding options for giving answers of a definite character. Other boundary conditions were that communication always remained possible and that there was a systematic way of enquiring about the philosophical question. Questions of a guessing nature, as 'who will be the winner of the next football game?' were excluded. Preferable were questions that appear somewhat unusual and difficult on the one hand, but with a powerful drive to ponder over as well.

Examples of such philosophical questions that do not really have an ultimate answer and can be discussed critically and systematically are:
- Do colours exist if no one can see them?
- Is a world champion sprinter really the fastest runner on Earth?
- Can a monkey that paints be an artist?
- If it would be possible to deep-freeze people, could they then live forever?
- Am I still myself when my body (or my name) would change?
- Why is the value of a renovated painting high and of a copy almost worthless?
- Can our senile granny still make decisions?
- Can reality be exchanged for virtual reality or a dream?

In the tetralogues, trains of thought were constructed in an attempt to provide insight into a philosophical problem.

Rules for discussion and duration

Before starting the session, youngsters were instructed that the given philosophical question would be discussed entirely among them. This included analysis of the question, search for evidence for their arguments, to reason and imagine counter arguments, and search for comparisons and possible answers. Some rules were set to stimulate production of autonomous and consistent thoughts undisturbed by external factors, and to avoid unlimited babbling and endless repetition of arguments. These rules concern linguistic and non-linguistic behaviour and aspects of communication.

Rules of the tetralogue

1. Opinions are only allowed when backed up by arguments.
2. Participants may attack arguments but not opinions of others.

3. Thoughts and arguments should preferably be consistent and constructed systematically; they should fit in the context of previous sayings, hypotheses and assumptions in the session.
4. Sayings and arguments must be accessible and controllable and must thus be in a language understandable to all and free from hidden presuppositions or introspection.
5. Dogmas, irrational certainties, arguments based on external authority or definitive judgements were not allowed since these are not considered rational arguments.

Unannounced interruptions were permitted encouraging expression of spontaneous ideas. Tetralogues started at the moment the philosophical question was clear to all. This process usually took between 5 and 15 minutes. Tetralogues typically ended when the school bell rang for another lesson or pause. Occasionally, the end of the audiotape – marked by a click-sound – terminated the tetralogue. In cases where tetralogues were conducted outside of school, the chair closed the session at an appropriate point in the discussion after approximately 45 minutes.

3 *Investigation into the psychometric qualities of the tetralogue*

Purpose of the investigation

The use of tetralogues as tools for assessing philosophical quality is partly validated in this section. To this end, the preconditions for construct validity will be checked against empirical results. The philosophical quality will be examined on two levels: 1) On the individual level (i.e., of each participant in a tetralogue); and 2) On the level of the whole discussion by the group. To this end, the quality of several lower level psychometric aspects of the tetralogue-performance will be investigated: objectivity of scoring, reliability of tetralogue performances, reliability of indicators, and consistency over time.

Methods, subjects, and selection of discussion groups

Discussions according to the tetralogue-format were used as research instrument. Data were collected on 112 tetralogues with 334 participants of varied backgrounds, in age, educational level and past biography. Some tetralogues were dropped because of logistic reasons. 95 tetralogues remained with 281 participants, distributed over 13 cells, designed in advance: 12 cells on behalf of the main project, and one cell with tetralogues as part of a longitudinal research (see Table 4.b). Each cell was filled with 5 to 10 tetralogues, with participants of approximately the same age and educational level. The additional longitudinal research was performed with the same four participants over a time span of 28 months. Two tetralogues of this longitudinal research were also used in the main project.

Variation between tetralogues is partly by design. On theoretical grounds, exposed in Chapter 2, age, educational level, and type of life course were expected to

influence the philosophical quality. Therefore, tetralogues data were collected according to a complete factorial [Age (3) by Educational level (2) by Life course (2)] design, shown in Table 4.b.

Three age categories were distinguished:
1. 11, 12-, and 13-years old
2. 14, 15, and 16-years old
3. 17, 18, and 19-years old.

The two categories of educational level were:
1. high or semi-high level (VWO and HAVO)
2. middle or low level (VMBO).

Searching for the influence of exceptional circumstances on the philosophical quality of thinking patterns, another two categories concerning the regularity of life course were introduced:
1. regular life course (regular schools)
2. irregular life course (physically disabled, imprisoned, or living without parents).

Table 4.b. Overall picture of all recorded tetralogues in cells of combined categories concerning age, educational level and life course.

Tetralogues of the main project			
Age: 11-12-13 years **Educational Level:** BO 8-HAVO-VWO 1-2 **Life course:** regular **N:** 10	**Age:** 11-12-13 years **Educational Level:** BO 8-VMBO 1-2 **Life course:** regular **N:** 8	**Age:** 11-12-13 years **Educational Level:** BO 8-HAVO-VWO 1-2 **Life course:** irregular **N:** 6	**Age:** 11-12-13 years **Educational Level:** BO 8-VMBO 1-2 **Life course:** irregular **N:** 5
Age: 14-15-16 years **Educational Level:** HAVO-VWO 3-4 **Life course:** regular **N:** 10	**Age:** 14-15-16 years **Educational Level:** VMBO 3-4 **Life course:** regular **N:** 7	**Age:** 14-15-16 years **Educational Level:** HAVO-VWO 3-4 **Life course:** irregular **N:** 8	**Age:** 14-15-16 years **Educational Level:** VMBO 3-4 **Life course:** irregular **N:** 9
Age: 17-18-19 years **Educational Level:** HAVO-VWO 4-5-6 **Life course:** regular **N:** 4	**Age:** 17-18-19 years **Educational Level:** VMBO 4-5-6 **Life course:** regular **N:** 4	**Age:** 17-18-19 years **Educational Level:** HAVO-VWO 4-5-6 **Life course:** irregular **N:** 9	**Age:** 17-18-19 years **Educational Level:** VMBO 4-5-6 **Life course:** irregular **N:** 8

Tetralogues purpose of longitudinal research
Time span: 28 months
Participants: Stefan, Martijn, Thomas, Erwin
Educational level: BO 8, VMBO 1, VMBO 2
Life course: regular
N: 9

Another source of tetralogue variation comes from the factor 'time'. One tetralogue discussion group was followed-up nine times over a period of 28 months.
Of course, inter-individual differences between participants are an important source of variation in philosophical performance. Variation in tetralogue and individual characteristics are investigated closer in Chapter 6.

Group composition

The demand of equality for group constitution only refers to the categories of age and educational level. Although regularity of life course is a criterion of selection, equality of group constitution in this respect is not a distinguishing feature in order to attain qualitative results in terms of a tetralogue. It was assumed that the level of philosophical quality of a tetralogue would increase when different life experiences play their part. This is the case when one, two, three or all participants have non-regular life courses. Sometimes it happens in a group that 'regular youngsters' carry very special life experiences, for example orphans, refugees or teenage-mothers.

This study also aimed at mixed genders in sessions, preferably with two boys and two girls. Unfortunately, there was sometimes very uneven gender distribution in specific groups (e.g., in a technical school or prison). 33 Tetralogues were composed of participants with the same gender.

In exceptional cases fewer or more than four youngsters participated in a session, for example if a participant forgot an appointment, or in cases of sudden illness or unforeseen visits to the dentist. In nine tetralogues (Bart 4, Alphen 2, Teyl 5, Teyl 8, Mk 4, Mk 5, Mk 8, Schot 3, and Schot 7) three students took part, and in Elis 1 and Schot 6, only two youngsters were present. These last two cases were part of a series of tetralogues with the same participants: a follow-up tetralogue and one with an incomplete group of adopted youngsters. In both cases, the dynamic of generating thinking patterns easily stand the demands of tetralogue. Sometimes a youngster not present in the first meeting appeared to be very eager to participate in the follow-up. Eight tetralogues had five participants instead of four (Bart 1, Bart 3, Dam 1, Dam 5, Dam 6, Everg 4, Everg 9, Wittevrouwen 8, Aml 2-3, Teyl 2, and Teyl 4).

In a few cases, not all physically present youngsters participated in the discussions. In a tetralogue performed in Brussels, one boy did not say a word: probably he did not understand the Dutch language. Eleven of the tetralogues had a participant that produced less than 10% of the utterances. This may have been due to shyness. Alternatively, participants may be noticeable interested, actively involved in the thinking process but refrain from orally expressing themselves. These cases may smoothly bridge between tetralogues with three participants. Recruitment of participants is discussed in more detail in Chapter 6.

Note with respect to the realisation of the design

With respect to the age of participants, it should be noted that physically handicapped participants of the same educational level were often older than their non-handicapped class-mates because of frequent absences. Other tetralogues had a not anticipated multiform character including participants with non-regular life course (e.g., one tetralogue included a teenaged mother and an orphan). These tetralogues were assigned to the category of 'irregular life course'.

Size of data set

In all, 112 tetralogues were recorded, but eight were rejected due to unintelligibility, inarticulateness, over-long introductory periods, technical shortcomings, or because too few students showed up. One tetralogue was rejected for ethical reasons and to protect the privacy of one or more participants. Eight tetralogues were not analysed because of a lack of time and they appeared to be over the cell limit (see Table 4.b). An overview is presented in Table 4.c. From one tetralogue part of the tape was lost, although the remainder was substantial and included in the research data set. Tetralogues with 17 to 19 year-olds have to be analysed yet.

Table 4.c. Diagram of tetralogues used, dropped out, and not yet analysed

Tetralogues	**Number**			**Number**
Recorded	112			
Rejected	17	Reason of rejection	unintelligibility	3
			technical and logistical shortcomings	5
			ethical reason	1
			over completion	8
Ready for use	95			
Not yet analysed	25			
Used for analyses	70			

Seventy tetralogues were accepted for analysis and used in this psychometric investigation. Some 216 youngsters participated, producing a total of 14, 393 utterances. Each utterance was scored for the presence of 6 binary philosophical indicators. In the following section, both individuals and tetralogues will be analysed. Examples of relevant data matrices are given in Tables 4.d and 4.e.

4 *Results concerning objectivity, reliability, and validity of the measurements*

Overview of reporting results

Scoring and preparation of the data for SPSS analyses will be reported in Section 4.4.1. The objectivity of scoring is analysed in Section 4.4.2. Reliabilities in its several forms are reported in Section 4.4.3. An index for individual's philosophical

quality is constructed in Section 4.4.4, and for the philosophical quality of a tetralogue as a whole in Section 4.4.5. Finally, some empirically based statements about the construct validity of the philosophical quality of individual contributions, and of a tetralogue as an instrument to assess philosophical quality, are made in Section 4.4.6.

4.1 From registration to *data*-processing

After a tetralogue was transcribed (see Section 4.2), utterances were assessed according to the presence of one or more of the six indicators. The process from registration of a tetralogue to assigning indicators to the youngster's utterances, and data processing had three distinct phases and was fully transparent and replicable. The first phase comprises registration, transcription and formatting in clear-cut utterances accessible for unambiguous interpretation. The second phase covered encoding of utterances by assigning indicators to the transcribed utterances. In the third phase, the encoded utterances were processed in statistical data files. While registration, transcription and processing of scores (encoded utterances) was achieved objectively, encoding of utterances included some judgements bound to strict rules to avoid subjective biases.

Registration

All tetralogues were both audio- and videotaped. An external microphone of the PHILIPS cassette recorder was placed on a table in the middle of the group and the VHS video camera outside the group at a distance of maximum three metres in a position securing sufficient light. The video camera and audio recorder were installed and recorded direct from the start of the tetralogue and were operated automatically. The presence of video camera and audio recorder did not seem to affect the participant's attitudes and had no observable impact on the discussions in the tetralogue. All video- and audiotapes were of the same brand. Only 90-minutes audiotapes and 3-hours VHS videotapes were used. Generally, two tetralogues were recorded on one audiotape, while up to three tetralogues are recorded per videotape. All tapes were systematically archived at the author's house and a catalogue is available on request.

Transcription

It is a fact of life that two observers sometimes report differently on the same event. Discrepancies in registered observation may be partly avoided by transcription. Therefore, all audio- and videotaped tetralogues were written-out (transcribed). This process was performed in two steps and mainly by students. First, the audiotapes were transcribed in full, followed by observing the videotape to personalise sayings (who said what), to check correctness of the sayings written from the audiotape, and to add non-linguistic expressions. Both the audio- and the video recorder were equipped with a provision to pause and play back the tapes. Texts were written

out digitally in Word. The full transcription process typically took some 8 hours per tetralogue.

Although extra-linguistic aspects such as intonation, voice-raising, body language, balance of power between participants experienced in a vivid and concrete situation do contribute to the meaning of the linguistic expression, these are not equivocal and probably reflect (semantically) nominal meanings more than discursive and argumentative thinking patterns. Extra-linguistic aspects in tetralogues are generally omitted since the aim of this study was confined to aspects of thinking patterns in which the person who carries them bears account, referring to meanings that can be depicted objectively. Exceptions are made for questioning glances, eyebrow movements, (questioning) intonations and shrugs. These attitudes may reflect qualities of Indecisive thinking (Idt), Openness (Op), Tentative behaviour (Te), and Epistemic position (Ep). Such extra-linguistic manifestations may be observed quite objectively and can thus be transcribed by means of a question mark (?). Extra-linguistic expressions as indecisive, doubtful, or problematic glances and body expressions coincide often with hesitations, stammering or flagging of the sayings. These expressions are transcribed by three dots (...). A last exception concerns the appearance of a clear-cut discrepancy between the nominal meaning of words used and the non-linguistic expression of a participant that must be noticed independently by two different observers involved in the transcription process. In these exceptional cases, this observation was reported in the transcription between brackets (...).

Another aspect of transcribing recorded observations is the use of punctuation. Without going into linguistic details, the meanings of two sentences separated by a full stop or not may be quite different. Even larger discrepancies may occur between sentences posed in the indicative way with a clear questioning intonation and inquiring glances and the same (indicative) sentence but written without a question mark.

Inaudible or inarticulate expressions are not transcribed. Broken but intelligible sentences are transcribed by ending with three dots (...). Meaningless words or expressions like 'eh', 'ha', etcetera, are literally transcribed as intelligible. Despite the preconditions set for the tetralogue, some youngsters were not in full command of the Dutch language. For the Netherlands, this seldom occurred in schools with a large allochtonous population, as foreign pupils generally did not volunteer to participate or kept very quiet. In Belgium, participants sometimes reverted to French vocabulary. Only when words and utterances were understood by all participants and chair, they were they encoded and transcribed. In other cases, they were ignored and denoted by means of dots in the transcription.

When a word was reiterated more than twice, the next and following iterations were denoted by three dots (...). In the rather exceptional cases of a long pause during a tetralogue the transcription displays ('stilte' = silence in Dutch). Often, sayings were interrupted by expressions of others. In such cases, the transcription of the first speaker's saying ends with three dots (...) and the interruption starts with a dash (-).

A transcription manual was composed to ensure uniformity in the transcription process. Despite of some residual lack of clarity, the written transcriptions constituted the foundation for assigning *philosphical quality* indicators to the utterances of youngsters in tetralogues.

Formatting

The author verified all transcribed tetralogues. Written texts were compared with the videotapes and corrected when necessary. Simultaneously, written texts were set in a special format to access the expressed utterances to the standardised encoding process during which *philosophical quality* indicators were assigned. This format consisted of a table containing two columns and many rows (Table 4.d). Each tetralogue is put into one table. In the table heading are indicated the location, the date, acronyms and names of participants, the chair, the duration, and the subject of the tetralogue. The transcribed expressions were rewritten in the right-hand column.

The column to the left contains three topics:
- Number of the utterance (unit of utterance)
- Acronym of the speaker
- Attributed *philosophical quality* indicator(s) (scores).

Each separate utterance occupies one row. A blank row is left in between two separate utterances reserved for various meaningless or unintelligible expressions ('background noise', see Table 4.d).

Table 4.d. Fragment from a tetralogue format about the possibility of an eternal life under perfect deepfreeze conditions with three female and one male participants, between 16 and 17 years old.

Elis 2	**41 min.**
Elisandra-groep 2	
Amersfoort, 24 augustus 2000	
E: Elisandra	
Mas: Macha	
Mar: Martiña	
Mo: Mouna	
T: chair, the author	
Do you live for ever when you are perfectly deep frozen?	
Introduction:	
Mo: I am Mouna. Mar: I am Martiña. Ma: Macha. E: I am Elisandra. T: And Elisandra? Who likes to tell ... Mouna. Do you like to explain the topic to discuss. Please try to eye each other.	
Ma: Well ... Mar: I don't even look at her! T: On which topic do you like to philosophise?	
Mo: ... whether you are able to deep freeze people. E: And if so, whether they do live forever.	
Ma: Do people live for ever when they ... are deep frozen. Is it correct like this? T: Yes. What do	

you think? But Martina just came up with rather interesting topic. Ma: Why? T: She continued on it directly when I said: stop, wait until the tape plays. What, what makes the topic interesting to you? What question did arise to you?

1 Mar Idt-op-re	Mar: Well the question you were stating about the situation of being deep frozen: if so, do you live for ever? But an eternal life, well, this you have only if you have lived, don't you? Well, if you have been deep frozen, then you will miss life, isn't it?
2 Mas Te-re	Ma: But I mean, if you breathe, I mean people that are in coma, they are still living, they are not dead, they live. If you are deep frozen, then ... yes ... then you are laying over there in your ice cubes ... -Mar: Yes, you are laying over there, but ... ... Yes, but you are breathing, your heart is still beating, so you are not dead, you are alive.
3 Mo Idt	Mo: There is no need for this.
	Mar: But then ...
4 Mas -	Ma: Yes, all right, they unfreeze you and you are dead, but then ...
5 Mar Op	Mar: But why is there no need for this?
6 Mo Te-ep	Mo: Well, they can ... yes, I think technology will ever evolve and come up with circumstances in which they are able to call you into being, or something like that.
7 Mas Op-te-ep	Ma: No, but if they deep freeze you and ... then ... they still take for granted that you keep being alive in your ice cube, and that you ... well you are all the same departing from that, I think, aren't you?
	E: Well, I don't know really. We are discussing something else, according to me ...
601	T: No, no. This topic we have been chosen. You have to try to continue this point. I am interesting in ... Mouna said: there is no need for this. Please explain.
8 Mo Idt-te	Mo: Yes, you will never know all things they will discover. I mean, they could just as well call you easily from dead into being or something like that. They may make your hart beating, or put an artificial heart in your body.
9 Mas Op	Ma: No, but if you are deep frozen in real, what about, then?

Unit of utterance

In each tetralogue, separate utterances were distinguished and described as units. An utterance is defined as an uninterrupted train of expressed thoughts or the continuous flow of sayings by one of the tetralogue's participants. Each utterance should therefore have a beginning and an end. It ends when another participant (or the chair) starts speaking-up in a way that influences the course of the discussion. This also marks the beginning of a next utterance. Inarticulate or meaningless interruptions by others are ignored when this would leave the sayings of the first speaker unaffected. All utterances are considered as units and numbered, ranging from 1 onwards. The lowest number of utterances recorded in a single tetralogue was 83, the largest 397 units. The duration of each utterance may differ widely. The mean number of units in the registered tetralogues of this study was 203, whereas the average duration of the operational part of a tetralogue was 37 minutes.

Interruptions and interventions

Any interruption affecting successive utterances was counted as a new utterance and received a successive number. When the first speaker returns to their earlier statement, a new number is given. Out of the muddle of sayings, contributions of the dominant speaker were noted and numbered. Utterances by the chair were termed interventions. Interventions on house/order keeping issues were noted as background noise and left unaccounted for. Content related interventions by the chair were numbered from 601 onwards. No indicators were assigned to such interventions.

Background noises

Apart from interventions by the chair, four types of sayings were considered as background noise and not accounted for. These utterances were filed in one (or more) intermediate rows: 1) Unintelligible sounds; 2) Not-seriously meant sayings; 3) Questions for repetition; and 4) Irrelevant meta-utterances.

Unintelligible sounds are all sounds that cannot be captured because these are too soft, too inarticulate, or were shouted. These include expressions in foreign languages (e.g., in Arabic) or in secret languages, and meaningless sounds not present in dictionaries. Not seriously meant sayings comprise all kinds of stopgaps, fillers and teasing comments. Questions for repetition like 'What did you say?' occur frequently. Repetitive utterances were considered single and counted as one. The same held for expressions like 'ja' (yes) when expressed only as a murmur of approval, 'eh', 'ha-ha', et cetera. Irrelevant meta-utterances were totally out of context and appeared when, for example, someone remembering an appointment with the dentist. This category also included asking for the time, permission to go to the toilet, or to discuss another topic.

In cases when it remained unclear if an utterance was seriously meant or if the speaker was joking or commenting as filler, the utterance was given the benefit

of the doubt and made assignable to indicators. Inevitably through the transcription and formatting process, many arbitrary decisions were made. In these cases, the availability of, and reference to the transcription manual proved very useful.

Assigning indicators to single utterances (encoding/scoring)

Once all utterances were identified and numbered, they were judged to see whether one or more philosophical quality indicators could be assigned to them. Judging utterances on their *philosophical quality* was a crucial component of this study.

More than one indicator could be assigned to one utterance, but an indicator may be assigned only once to an utterance, even if this indicator appeared more often in the total utterance. So, the number of indicators assigned to one utterance ranged from 0 to 6. Initials of the indicator(s) were placed in front of the numbered utterance in the left column of the table (Table 4.d). An extensive register of expressions (in Dutch) characteristic for each indicator is attached as Appendix II. This list proved to be particularly helpful when determining if more than one indicator occurred in an utterance. Scoring is a qualitative process and may not be free from personal bias. Objectivity and reliability of this scoring process are discussed in Sections 4.4.2 and 4.4.3 respectively.

Most problems in the scoring process were encountered in two types of situations:
1. Expressions assignable to more than one indicator depending on its meaning.
2. Unintelligible or ambiguous remarks.

Expressions in which two compatible indicators may be distinguished were assigned double. Examples are given in the register. This particularly applied to rhetorical questions assigned to both Openness (Op) and reasoning quality (Re); 'I do not know', assigned to Indecisive thinking (Idt) and as Epistemic position (Ep); and 'if-then' sentences, starting with a contra factional hypothesis, assigned to Tentative behaviour (Te) and Reasoning quality (Re). The same held for most analogies because they are illustrations and trials to joint reasoning. Pushing statements with a wondering overtone were assigned to Openness (Op) and Tentative behaviour (Te).

Expressions where two or more incompatible indicators occur were ambiguous to judge and difficult to assign indicators to: assignment to one indicator would consequently exclude the other. The use of the word 'crazy', 'odd', or 'strange' for instance, mostly expresses a feeling of uneasiness and should thus be assigned to Indecisive thinking (Idt). But in rare cases a kind of openness is demonstrated that is not compatible with uneasiness. Expressions of doubt may also indicate either a sense for ambiguity (Indecisive thinking) or a way to express a trial (Tentative behaviour). But these two indicators cannot be valid simultaneously in this expression. In expressions like 'it depends' the sense may be of an ambiguous character: 'it may be this way, it may be that way'. But it may also indicate a clear network in which the concept is hanging. While in case of the first meaning, Idt has to be assigned; the last is rather a case of analysing and reasoning.

Unintelligible or ambiguous remarks are judged with nominal meanings and primarily seen independently of previous remarks. Often these are low profile and seem to be neither of any interest nor relevance with respect to philosophical qualities. However, they also may have a clear *philosophical quality* appearance but are actually used as fillers. For example, the use of the Dutch word 'eigenlijk', which has a meaning close to 'really', 'in fact', or 'honestly spoken' (see description Indicisive thinking in Section 4.1). Judging utterances with this kind of expression presuppose investigation in the linguistic performance and preceding utterances of a speaker. To judge if a speaker uses a word in a meaningful way or as filler, context, previous statements, and the evolution of the tetralogue was noticed.

Philosophical quality indicators cannot be assigned to interventions by the chair and to background noise as denoted in the rows between separate utterances. Nor can they be assigned to inaudible and inarticulate words, (parts of) sentences, and reiterations without extra information, fillers, stopgaps, teasing comments, or to any other sound registered during an utterance. This also holds for bare definitions, explanations, descriptions and quotations. Question marks are assigned to Openness (Op). If the mark of three dots appears somewhere in an utterance (not at the end) it indicates hesitation and should then be assigned to Tentative behaviour (Te). Sometimes meaningful words are used as expression of caution or embarrassment and could not be assigned to an indicator. This happened often when a participant started with 'according to me', 'indeed' or 'it seems to me'. These expressions were used here as fillers. Another problem deals with the use of causal conjunctions meant as temporal conjunctions and the use of 'because' starting a clear-cut answer. Children often sum-up occurrences with 'because' or 'for' and begin their answers with 'because': 'I raised because the alarm rang' and a girl answers the question: 'Why are you happy?' 'Well, because I am a girl'. Such usage of 'because' cannot be considered as indicating Reasoning quality (Re) and therefore was not be assigned to that indicator.

Notwithstanding the many vague expressions of thinking patterns, judges may also encounter strong confirmations of their judgement. Often, other expressions in the same utterance support assignments to a specific indicator. Judging utterances on their philosophical quality by assigning *philosophical quality* indicators cannot be done without preparation. For this purpose, judges need to be trained, a process that takes at least some hours, in addition to studying the register.

PQ and pq

Determining the philosophical quality of utterances may not only refer to the philosophical quality of individual youngsters in a tetralogue; their accumulative effect may also highlight the philosophical quality of a tetralogue (PQ). The philosophical quality of a youngster's contribution to a tetralogue (pq) is a weighted sum of the indicators derived from his utterances in the tetralogue. The method to determine pq, as the individual philosophical quality expressed in a tetralogue, is described in Section 4.4.4.

The philosophical quality of an entire tetralogue (PQ) is also calculated on the basis of three elements: 1) Individual pq's; 2) Dialogical effects; and 3) The proportionate valued part of the tetralogue. The value of each tetralogue was also estimated independently on a [0-5] point scale. This estimation, performed in advance, after seeing the video-recording of the tetralogue only, is based on five parameters, traced back from the definition, presuppositions and main features of philosophy (see Chapter 2): a) The production of autonomous thoughts by participants; b) the intellectual distance to the philosophical question; c) The ability to generate new questions, wonderings and uncertainties; d) The understanding of the complexity of the matter; and e) The ability to persist in systematic searching. A full description of the calculations of pq and PQ are given in Sections 4.4.4 and 4.4.5.

Data Processing

All data needed for statistical analyses are presented in the left-hand column of tetralogue format (Table 4.d). These data comprise the (acronyms of) participants, their (numbered < 600) uttered thoughts, the numbered interventions (> 600), and the assigned indicators of philosophical quality to each of these utterances. These values are imported into a statistical programme (SPSS). Numbered utterances and interventions were plotted along the vertical axis and indicators (maximum 6) assigned to utterances of all of the individual participants (normally four) plotted along the horizontal axis of the spreadsheet (Table 4.e). The cells contain in each horizontal row (i.e., each utterance or turn of a participant), values representing the presence of the pq indicators for each participant. Three values can be distinguished:

0: no utterance made by participant, no indicator assigned;
1: utterance made by participant, but an indicator is not assigned;
2: utterance made by participant and the respective indicator is assigned.

In total 14, 393 utterances by 216 youngsters participating in 70 tetralogues were assessed regarding their philosophical quality by scoring the presence of the six indicators in each utterance. In this way, each utterance consists of six binary components, and for each participant, all utterances were scored (Table 4.e).

Two other data files were compiled. One (participant file) consists of personal data for all individual participants, including tetralogue numbers, age, gender, educational level, level of regular life course, GPA (Grade Point Average) languages, results on the evaluation questionnaire, results on the Raven test, results on the Neo FFI (see Chapter 6), assigned indicators, philosophical topic, number of utterances made personally, total number of utterances in the tetralogue, duration of the tetralogue, and computed philosophical quality. The other file (tetralogues file) consists of data for all tetralogues, including: tetralogue numbers, average age of participants, educational level, level of regular life course, philosophical topic, number of utterances per tetralogue, number of interventions, duration, chairperson, number of male participants, number of female participants, estimated philosophical quality, and calculated philosophical quality.

Table 4.e. Example of a part of SPSS data file for the tetralogue Elis 2. On the vertical (Y) axis: the turn, i.e. the serial number of the utterance. On the horizontal (X) axis: the presence of each indicator (i, o, t, e, r, a) for each participant (m, e, mt, mo).

turn	mi	mo	mt	me	mr	ma	ei	eo	et	ee	er	ea	mti	mto	mtt	mte	mtr	mta	moi	moo	mot	moe	mor	moa	i
1	0	0	0	0	0	0	0	0	0	0	0	0	2	2	1	1	2	1	0	0	0	0	0	0	0
2	1	1	2	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	1	1	1	1	0
4	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	1	2	1	1	1	1	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	2	1	1	0
7	1	2	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
601	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	2	1	1	1	0
9	1	2	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Legend: 0: no utterance made by participant, no indicator assigned

1: utterance made by participant, but an indicator is not assigned

2: utterance made by participant and the respective indicator is assigned

Missing values

In less than 2% of the tetralogues the age of participants was not communicated to the author. In these cases, missing ages were substituted by the average age of the other participants in the tetralogue. This was since belonging to the same age group was a pre-condition on participant selection. Missing information on education level was more difficult to substitute. This was particularly true for mixed tetralogues of participants with irregular life course. This was the case for a tetralogue with three 18-year old participants in the closed penitentiary institution of Teylingereind. At some point in the past, these students had interrupted or left school, so their education level was attributed a lower level.

4.2 Objectivity of scoring, estimated by inter-rater agreement between judges

Assigning indicators to utterances is a qualitative activity and for the purpose of this study was made transparent, objective and reliable. This section concerns objectivity. Objectivity of scoring refers to 'declaring facts'. The plausibility of these 'observed facts' requires that at least two independent observers or judges come to the same judgement. Objectivity in judging utterances is tested by a. the proportion of raw agreement between two different judges of the same tetralogue, and b. by Cohen 's Kappa, a coefficient for inter- and intra-rater agreement that corrects for chance-agreement (Cohen, 1968; Bakeman & Gottman, 1997).

Judges

Qualitative judgements of utterances cannot be mechanised but must be performed by a person – a judge. The judge should identify *philosophical quality* indicators in thinking patterns based on carefully transcribed uttered expressions by youngsters in tetralogues. Identification of such indicators was based on recognising constitutions of utterances, semantically, syntactically and practically, in the perspective of a developing tetralogue and of cultural conventions of participants and of the judge's own. The outcomes of this complex process were described in terms of scoring results: a certain indicator was present or not present. These scores were made as objective as possible, with objectivity being reflected by the level of agreement amongst the judges.

To this end, several co-judges were invited to independently evaluate philosophical qualities of youngsters through the indicators assigned to utterances in tetralogues. Since it is not possible to understand one utterance separately from its context, each co-judge was invited to assign indicators to all utterances in a complete tetralogue. Instead of looking at all six indicators co-judges were asked to focus on five and ignore utterances with an Anecdotal quality (An) because this indicator is not vulnerable to bias as it presence is scored only when references are made to places, time or circumstances reported in utterances (see Section 4.1). All co-judges were instructed by the author and provided with the register (Appendix II). Two students in philosophy who had conducted a tetralogue previously, were asked to evaluate these tetralogues. In other cases, the co-judges received randomly selected tetralogues.

Two co-judges were students in Philosophy. These students were the only co-judges with past experience of conducting tetralogues. Instruction consisted of a group training session by the author, with the register and a personal oral clarification. Because their judging activities were assessed in an early phase of the study, not all indicators were fully described. Therefore, their scoring results are comparable with those of the author only to a limited extent. Neither were scoring results concerning different tetralogues mutually comparable because of their different contexts. The co-judges performed their activities prior to review of the tetralogues by the author. One co-judge was an outsider in philosophical and psychological matters, and inexperienced with the problems of philosophical discussions. This judge received the same instructions as the Philosophy students and then evaluated two full tetralogues together with the author, then three tetralogues were judged independently. One co-judge was a philosophically educated psychologist and an experienced judge who operated with the instructions in the register and with some oral clarification.

Inter-rater agreement was assessed by comparing the scoring results of the author (D*) and four co-judges (A1, A2, B, and C). One co-judge (B) evaluated three tetralogues (B1, B2, and B3). All scoring results were processed in the statistical programme SPSS, from which the proportion of agreement was produced. The inter-rater coefficient, Cohen's Kappa, correcting this agreement for chance-agreement, was also calculated. Also, intra-rater coefficients were calculated by

comparing the scoring results of five tetralogues (D1, D2, D3, D4, D5) with D*. All these tetralogues were judged by the author at two different times more than six months apart. When agreement was found between judgements on assignment of indicators to utterances, the presence of a common underlying factor to both parameters was assumed likely. The percentages of agreement and the values of Cohen's Kappa on the scoring results by the co-judges and the author for five indicators are given in Table 4.f, 4.g, 4.h, 4.i, 4.j.

Table 4.f. Inter-rater agreement (between judge A, B, C and D*) and intra-rater agreement between D1 to D5 and D* for **Indecisive thinking (Idt) indicator**.

Second Judge	Name of the tetralogue	Total number of scored utterances	Percent of agreement	Value of Cohen's Kappa
A1	Stedelijk 1-1	245	82.0	0.63
A2	Dinxperlo	125	76.0	0.45
B1	Schot 1	263	86.3	0.56
B2	Jordan 1	114	80.7	0.56
B3	Elisandra 2	171	86.6	0.62
C	Hengelo 3	274	86.5	0.67
D1	Schot 1	263	94.4	0.80
D2	Stedelijk 3-3	181	97.3	0.91
D3	Hengelo 4	211	93.8	0.83
D4	Damiate 5	329	93.3	0.69
D5	Elisandra 2	171	94.1	0.82

Table 4.g. Inter-rater agreement (between judge A, B, C and D*) and intra-rater agreement between D1 to D5 and D* for **Openness (Op) indicator**.

Second Judge	Name of the tetralogue	Total number of scored utterances	Percent of agreement	Value of Cohen's Kappa
A1	Stedelijk 1-1	245	79.0	0.30
A2	Dinxperlo	125	93.6	0.47
B1	Schot 1	263	94.6	0.69
B2	Jordan 1	114	88.6	0.64
B3	Elisandra 2	171	90.7	0.72
C	Hengelo 3	274	95.2	0.80
D1	Schot 1	263	97.0	0.83
D2	Stedelijk 3-3	181	98.3	0.94
D3	Hengelo 4	211	96.7	0.91
D4	Damiate 5	329	95.4	0.76
D5	Elisandra 2	171	95.9	0.87

Table 4.h. *Inter-rater agreement (between judge A, B, C and D*) and intra-rater agreement between D1 to D5 and D* for* **Tentative behaviour (Te) indicator**.

Second Judge	Name of the tetralogue	Total number of scored utterances	Percent of agreement	Value of Cohen's Kappa
A1	Stedelijk 1-1	245	66.0	0.27
A2	Dinxperlo	125	88.8	0.78
B1	Schot 1	263	93.5	0.63
B2	Jordan 1	114	81.6	0.62
B3	Elisandra 2	171	82.5	0.64
C	Hengelo 3	274	93.4	0.86
D1	Schot 1	263	89.7	0.79
D2	Stedelijk 3-3	181	97.8	0.96
D3	Hengelo 4	211	91.9	0.83
D4	Damiate 5	329	92.4	0.83
D5	Elisandra 2	171	89.5	0.79

Table 4.i. *Inter-rater agreement (between judge A, B, C and D*) and intra-rater agreement between D1 to D5 and D* for* **Epistemic position (Ep) indicator**.

Second Judge	Name of the tetralogue	Total number of scored utterances	Percent of agreement	Value of Cohen's Kappa
A1	Stedelijk 1-1	245	95.0	0.88
A2	Dinxperlo	125	78.4	0.53
B1	Schot 1	263	93.5	0.80
B2	Jordan 1	114	95.7	0.91
B3	Elisandra 2	171	92.4	0.84
C	Hengelo 3	274	97.8	0.95
D1	Schot 1	263	98.4	0.96
D2	Stedelijk 3-3	181	95.0	0.88
D3	Hengelo 4	211	97.6	0.95
D4	Damiate 5	329	98.2	0.92
D5	Elisandra 2	171	99.4	0.99

Table 4.j. *Inter-rater agreement (between judge A, B, C and D*) and intra-rater agreement between D1 to D5 and D* for* **Reasoning quality (Re) indicator**.

Second Judge	Name of the tetralogue	Total number of scored utterances	Percent of agreement	Value of Cohen's Kappa
A1	Stedelijk 1-1	245	82.0	0.64
A2	Dinxperlo	125	90.4	0.45
B1	Schot 1	263	82.9	0.48
B2	Jordan 1	114	82.4	0.50
B3	Elisandra 2	171	78.4	0.29
C	Hengelo 3	274	93.4	0.81
D1	Schot 1	263	92.0	0.79
D2	Stedelijk 3-3	181	97.3	0.93
D3	Hengelo 4	211	97.2	0.92
D4	Damiate 5	329	93.9	0.84
D5	Elisandra 2	171	90.6	0.71

If the scoring results on Indecisive thinking (Idt), Openness (Op) and Tentative behaviour (Te) by student judges (A1 and A2) are discarded, then the Tables 4.e, 4.f, 4.g, 4.h and 4.i show that the percentage of agreement for all indicators exceeds 78.4, for the judgement of the reasoning quality at B3. The values of Cohen's Kappa on the judgement of reasoning quality by the same inexperienced co-judge are low: 0.48, 0.50 and 0.29. Other values of Cohen's Kappa are all higher than 0.56 (also found at B). These figures demonstrate that the categories for the qualitative judging process of assigning indicators to registered utterances of youngsters can be scored objectively at least when the scoring is done by experienced raters.

Most of the discrepancies between the scoring results by the author and co-judges were attributable to failures rather than differences in interpretation on assigning indicators or where a co-judge misses noticing the indicator in an expression. These results also demonstrate differences in the level of development of indicators, in judging different indicators, in backgrounds and degree of experience of co-judges, and in learning effect during different times. In an early stage of indicator development, percentages of agreement and values of Cohen's Kappa appeared on a clear lower level, albeit with an approximate significance of 0.00. In all cases the Epistemic position (Ep) appears to be relatively easy to identify reliably. In the potential 9 cases of comparison in which a matured level of indicator development is operating, Cohen's Kappa varies from 0.80 to 0.99. Indecisive thinking (Idt) was the most difficult indicator to identify in the registered utterances of youngsters. In 4 of the potential 9 cases, the value of Cohen's Kappa was higher than 0.8 due to the intra-rater coefficients. Identifying a reasoning quality seemed to be difficult only to the inexperienced co-judge.

Regarding the co-judges' background, it must be noted that with the exception of judge B, all were philosophically educated. In spite of some trials, students of psychology could not successfully complete the judging task. The assessment process showed that those with a psychology background are better chairs, while the scoring results of those with a philosophical background are compare favourably with the results of the author. Experience matters as can be noted from the high degrees of agreement in the scoring results of co-judges C and D. After considerable training, inexperienced co-judge C reached an acceptable degree of agreement but was unable to increase this level of agreement over time. There was a time lag of two months between the judgements B1 and B2, and also between B2 and B3. However, no learning factor in this judging process can be demonstrated. Learning and experience must be distinguished. During the judging process the level of experience is supposed to be constant, although learning effects may be derived from remembrance and recognition of meanings. With respect to the comparison of judgements of the author (intra-rater), remembrance may play a role, but there was a limit on the time available to remember 250 complicated utterances. As the unchangingly high values of the intra-rater coefficients for judge D demonstrate, the high level of experience seems to guarantee consistent scoring.

4.3 Reliability

In Section 4.4.2 it was demonstrated that assigning philosophical quality indicators to registered utterances of youngsters in tetralogues is a qualitative, but essentially replicable process. This suggests that indicators are insensitive to subjective rater biases. However, the production by the individual of utterances in a tetralogue might be a largely random process. This means that the consistency or reliability of the utterances must be demonstrated. Reliability will be investigated on two levels: for the tetralogue as a whole and for each indicator separately.

4.3.1 Reliability of tetralogue

The internal consistencies of the tetralogues were analysed by comparing two equivalent halves of each tetralogue. Evenly numbered utterance sets were compared with the corresponding unevenly ones. For each tetralogue, data sets exist of dichotomous scoring results for each participant's performance on each of the 6 PQ indicators (columns, usually = 24) with respect to each of the numbered utterances that generally varied between 100 and 300. Seventy tetralogues were available for analysis; most of them are performed by four participants, but a few were performed by three or five participants. The distribution of the tetralogues according to age group, level of education, and of regular versus irregular life course is shown in Table 4.b.

Analysis procedure

As any discussion will show some drift or evolution, by splitting a discussion into two halves, this drift or evolution should be equally expressed in both halves. When participants in two equivalent halves of a tetralogue show identical performances, it may be concluded that as tetralogue has a high reliability (i.e., it is relatively insensitive to random fluctuations). To prevent disturbance from drift, utterances in tetralogues were attributed to evenly numbered (E-turns) and unevenly numbered ones (UE-turns). When a tetralogue ended in an uneven utterance, this utterance was discarded, as were interventions by the chair. Each tetralogue data-matrix was re-arranged so that the evenly numbered utterances formed the first part of the rows and the unevenly numbered the second. Then the data-matrix was transposed: i.e., rows and columns were interchanged, resulting in a matrix of 24 rows (4 participants times 6 indicators), and a number of columns corresponding with the groups of evenly and unevenly numbered utterances. For each participant-indicator combination, a total score was calculated for the even and uneven halves of the tetralogue (the number of times a participant scored that particular indicator). The total score was the sum of the (dichotomous) specific indicator-occurrences over all the relevant utterances (see Table 4.k).

As a first step in the estimation of the reliability of the tetralogue, a correlation was calculated between the total scores on the even and uneven halves of the tetralogue. Each participant's total indicator combination was considered a replication. In the second step, correlation coefficients were corrected according to the Spearman-Brown formula to transform the estimate of the reliability of a half tetralogue into

an estimate for the tetralogue in its entirety. This correction was performed for each tetralogue. The results for the group of 70 tetralogues are shown in Figure 4.1 and Table 4.l.

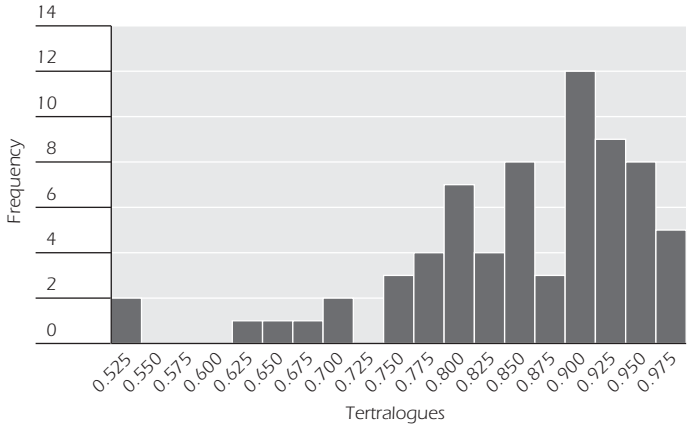

Figure 4.1. Distribution of split-half reliabilities (corrected cf. Spearman-Brown) for 70 tetralogues.

Table 4.k. Illustration of the transposed SPSS data-matrix of one tetralogue with four participants (A, B, C, and D) demonstrating the appearance of six indicators in a total of n utterances (n = the last even number of utterances).

		Even numbered turns			Uneven numbered turns			Total even	Total uneven
		1	**3**	**n-1**	**2**	**4**	**n**	**turns**	**turns**
A	Idt	0	0	0	2	0	0	9	12
	Op	0	0	0	1	0	0	0	3
	Te	0	0	0	2	0	0	13	13
	Ep	0	0	0	2	0	0	10	12
	Re	0	0	0	1	0	0	3	5
	An	0	0	0	2	0	0	2	2
B	Idt	1	0	0	0	0	0	5	4
	Op	1	0	0	0	0	0	1	2
	Te	2	0	0	0	0	0	3	5
	Ep	2	0	0	0	0	0	5	3
	Re	1	0	0	0	0	0	0	2
	An	1	0	0	0	0	0	2	1
C	Idt	0	0	0	0	2	1	9	12
	Op	0	0	0	0	2	1	1	4
	Te	0	0	0	0	2	1	8	11
	Ep	0	0	0	0	2	2	8	10
	Re	0	0	0	0	1	1	0	2
	An	0	0	0	0	1	1	1	0
D	Idt	0	2	1	0	0	0	15	6
	Op	0	2	1	0	0	0	3	2
	Te	0	2	1	0	0	0	17	11
	Ep	0	2	2	0	0	0	18	6
	Re	0	1	1	0	0	0	2	3
	An	0	1	1	0	0	0	2	2

Table 4.l. Split-half reliabilities (corrected cf. Spearman-Brown) for 70 tetralogues.

N:		70
Mean:		0.85
Median:		0.88
Mode:		0.89
Std. Deviation:		0.10
Minimum:		0.52
Maximum:		0.98
Percentiles	25:	0.90
	50:	0.88
	75:	0.93

Figure 4.1 shows the distribution of the varying split-half reliabilities. The mean reliability of the 70 tetralogues was 0.85. However, two tetralogues had low reliabilities (0.53 and 0.54), of which one was part of a series of 9 tetralogues with the same participants, executed under similar circumstances showing higher and acceptable reliabilities (0.80, 0.71, 0.93, 0.87, 0.91, 0.90, 0.78). Another tetralogue developed into a debate between two youngsters who sharpened their arguments against one another. In this case, the division into even and uneven turns was probably not a division into equivalent halves of the tetralogue. Nevertheless, the reliability of 75% of the tetralogues exceeds 0.80 for this study.

According to Nunnaly & Bernstein (1994) reliabilities have to exceed 0.7 for good instruments. For this study, only three tetralogues showed reliabilities between 0.6 and 0.7 (besides the two mentioned). The relatively low number of utterances in these tetralogues is striking: 79, 86 and 124; the mean number of utterances in a tetralogue is 203.

To verify that these reliabilities were independent of the type of philosophical question discussed, a one-way ANOVA (analysis of variance) was executed, comparing five groups of tetralogues each with a similar philosophical question. These questions will be discussed in detail in Chapter 5, but are presented in outline here:
1. Epistemological and metaphysical questions
2. Anthropological questions
3. Ethical questions
4. Questions of concept analysis, meaning and demarcation
5. Others.

The results of this ANOVA analysis are shown in Table 4.m.

Table 4.m. Results of ANOVA analysis in which five types of tetralogues differing in topic are compared concerning their reliabilities.

ANOVA Topic	**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Between Groups	2,743	6	0.457	0.629	0.707
Within Groups	46,552	64	0.727		
Total	**49,296**	**70**			

No significant difference in mean reliability between types of philosophical questions was found. Independence of the reliabilities from types of philosophical questions may thus be concluded.

Independency of tetralogue reliability from participant's age, regular or non-regular life course and educational level were tested similarly. Again, no significant differences in mean reliability between groups composed of participants with similar traits were found concerning age, educational level, and life experience (Table 4.n).

Table 4.n. Results of ANOVA analyis in which Age (2), Educational level (2), Life course (2) are tested on their relation with tetralogue reliability.

Tests of Between-Subjects Effects					
Dependent Variable: R4LOGUE					
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Corrected Model	2.567E–02 [a]	7	3.667E–03	0.350	0.927
Intercept	46.691	1	46.691	4457.712	0.000
AGE	9.608E–03	1	9.608E–03	0.917	0.342
EDUCA	1.598E–07	1	1.598E–07	0.000	0.997
EXPERIEN	1.042E–03	1	1.042E–03	0.099	0.754
AGE * EDUCA	2.239E–03	1	2.239E–03	0.214	0.645
AGE * EXPERIEN	9.839E–04	1	9.839E–04	0.094	0.760
AGE * EDUCA * EXPERIEN	6.580E–03	1	6.580E–03	0.628	0.431
Error	0.649	62	1.047E–02		
Total	**51.502**	**70**			
Corrected Total	**0.675**	**69**			

a. R Squared = 0.038 (adjusted R Squared = –0.071)

It may thus be concluded that equivalent halves of tetralogues show consistent results with reliabilities in 75% of the tetralogues, exceeding 0.70. These reliabilities are not influenced by tetralogue's topic, and also not by age, educational level, and life course of individual participants.
Now, the focus will turn to the reliabilities of the separate indicators.

4.3.2 Reliabilities of separate indicators

Reliabilities of the 6 indicators are estimated in two ways: first, by the split-half method, and secondly, by inspecting replication of indicator performance in a subgroup of participants over two tetralogues (test-retest reliability).

Analysis procedure

Split-half method

Re-arranged tetralogues (see Section 4.4.2) were placed in a new data-matrix producing a matrix of 70 (the number of tetralogues) times 24 (4 participants ×

6 indicators) rows. The total scores for the even and uneven half of the indicator-occurrences in each tetralogue were placed in columns. For the indicators Idt, Op, Te, Ep and Re, the Pearson correlation was calculated between the sums of their even and uneven occurrences (N: 265 replications). The single occurrence of indicator An has no philosophical significance to individual performances and was not considered in this analysis. The results are shown in Table 4.o. Reliability, r(c), was assessed after correction according to Spearman-Brown (Drenth, 1990, p. 122).

Table 4.o. *Split-half reliabilities of indicator performances (r), and their Spearman-Brown corrected counterparts (r(c)).*

	r	**r (c)**
Indecisive thinking (Idt):	0.69**	0.82
Openess (Op):	0.71**	0.83
Tentative behaviour (Te):	0.70**	0.83
Epistemic position (Ep):	0.77**	0.87
Reasoning quality (Re):	0.73**	0.84

** Sig.: $p < 0.001$

It may thus be concluded that indicator frequency has a good reliability; it shows that random factors do not substantially influence the tetralogue's total indicator scores.

Replicability of indicator frequency in the same participants over two different tetralogues

Next, the contributions of participants in more than one tetralogue were compared. Eighteen participants acted in two tetralogues, seven in three, and two in four (N: 27). From the last nine participants, only their contributions to the first and second tetralogues were used. So, test-retest correlations were calculated between two consecutive tetralogues for 27 participants. Because the intended correlations concern different tetralogues of different durations and containing a different number of utterances, the total scores representing participant's contributions to the tetralogues have to be made comparable with each other. This was done by a correction for duration and for the number of utterances in each tetralogue. All total scores of indicator occurrences of one participant concerning a particular tetralogue were divided by its duration, expressed in minutes, and by its number of utterances. The outcomes for each relevant indicator concerning the first tetralogue were compared with those of the second. The results are shown in Table 4.p.

Table 4.p. *Test-retest reliabilities of indicator performances of the same participants over two tetralogues based on 27 replications.*

	r
Indecisive thinking (Idt):	0.42*
Openness (Op):	0.17 n.s.
Tentative behaviour (Te):	0.56**
Epistemic position (Ep):	0.62**
Reasoning quality (Re):	0.51**

* Sig.: $p < 0.05$ ** Sig.: $p < 0.01$; n.s.: non-significant

With the exception of Op, these figures representing correlations between indicator performances over two tetralogues demonstrate the stability of individual differences in indicator production over time. The intervals between the two tetralogues ranged from one day (4 participants), 7 days (4 participants), 10 days (2 participants), 14 days (3 participants), 28 days (1 participant), 41 days (2 participants), 43 days (4 participants), 77 days (4 participants) and 99 days (1 participant), while 2 participants contributed to two tetralogues on the same day. With respect to openness, the correlation is low and not significant. According to the preceding results, all indicators seem to be insensitive to random fluctuations. Low correlations between different tetralogues may be the consequence of several factors, including differences in the composition of participants and in philosophical themes, learning effects, or a missing linear relation between two consecutive tetralogues. Differences between philosophical themes will be taken up in Chapter 5. Differences between participants will be examined in Chapter 6.

4.4 Constructing an index for philosophical quality of an individual's contribution to a tetralogue

It was shown in Section 4.4.2 that indicators can be objectively assigned to utterances in a tetralogue. In Section 4.4.3 the internal consistency of a tetralogue was demonstrated, both as a whole and on indicator level. Also, stability of indicator performance over time was demonstrated for four out of five indicators. Another measure of reliability of the set of indicators is Cronbach's coefficient alpha, which also indicates the degree of homogeneity or internal consistency of this set. Internal consistency, computed by Cronbach's alpha, resulted in the value $\alpha = 0.80$ for the five indicators. This value is good (Nunnaly & Bernstein, 1994). This value of alpha signals the plausibility of a common factor underlying the utterances of single individuals. Therefore, it makes sense to construct a single index representing the philosophical quality of an individual's performance in a tetralogue.

This section will describe the construction of such a numerical index (pq). A numerical index representing the philosophical quality of a tetralogue as a whole (PQ) will be presented in Section 4.4.5. The status of a numerical index will be investigated as the expression of an assumed underlying factor. Individual pq indices make it possible to compare philosophical qualities performed in tetralogues by different individuals.

Calculation of pq

An individual's pq is derived from the indicators assigned to utterances. Each participant contributing in a tetralogue performs a number of utterances to which one or more indicators may be assigned. In different settings, the indicators Idt, Op, Te, Ep, and Re are responsible for philosophical quality of an individual's performance in a tetralogue. The way indicator scores reflect aspects of a philosophical quality contrasts with the simple interpretation of indicators outlined

in Section 4.1. Some indicators are more valuable than others, reflecting a higher degree of philosophical quality. Combinations of indicators in one utterance may be more valuable than the sum of such indicators assigned to separate utterances.

A pq index of an individual's performance in a tetralogue is based on participant's contributions and on frequencies of indicators. The pq index must also reckon with combinations of indicators that reflect a surplus value, while other singular or unpaired occurrences of indicator performances have little philosophical quality. Some combinations of indicator performances offer a surplus value. Performances expressing Idt and Op indicate a sense for ambiguity, intellectual uncertainty, vagueness, relativity, openness, and wondering. Idt and Op occurrences realise comparable aspects of pq. The value of Idt and Op occurrences increases when the performance of these indicators is combined with the performance of reasoning or tentative behaviour in the same utterance. Reasoning indicates systematic searching in a constructive way to build up thinking patterns. Tentative behaviour mostly indicates a kind of risky, speculating thinking and a vulnerable attitude also reflecting a kind of openness. So, combinations of Idt or Op with Re or Te create a surplus value to pq. Combinations of Te & Re often indicate an 'if-then'-relationship or analogical reasoning. Both represent constructive thinking. Combinations of Idt, Te & Re or Op, Te & Re do so likewise and generate surplus value. Performances in which the indicator Ep combined with Re mostly offered points of view stocked with reasons why and an autonomous production of a constructive line of thoughts.

Some singular performances are relatively valueless. An epistemic position is claimed as a standard for reflection, indicating deliberative considerations about fundamental questions. However it often happens that this indicator is not accompanied with questions, signs of ambiguity, tentative or reasoning behaviour and only serves as an expression of private opinion without arguments or reasons. Reflective ways of speaking, like 'I think that,' indicates a predetermined stance instead of an expression of open mental susceptibility, constructing an autonomous line of thoughts. Often it is also an expression of prudence. In all those cases the singular performance of an epistemic position in one utterance does not fulfil pq. Therefore, a correction is needed for this abuse of opportunity.

Likewise, singular performances of reasoning or tentative behaviour do not necessarily load the philosophical quality. The performance of a Te score sometimes indicates a lack of cooperation in the construction of lines of thought. Appearances of Te, Ep and Re carry some potential to influence philosophical quality when in combination with other indicators. A single performance of an anecdotal quality does not influence the pq index. This sixth indicator was developed to test for a possible correlation between the appearance of an anecdote and the performance of regressive abstraction over the entire tetralogue, illustrating an idea of the non-tangible world.

An individual's pq is always based on their performance within utterances and proportionate contribution, and can be derived as follows:

A pq index is the sum of the frequencies of five indicator occurrences (Idt, Op, Te, Ep, Re), *plus* the sum of the surplus valued combinations (Idt & Te, Idt & Re, Op & Te, Op & Re, Te & Re, Ep & Re, Idt & Te & Re, Op & Te & Re) of indicator occurrences, *minus* the sum of the frequencies of unpaired occurrences of Te, Ep or Re; this total must be divided by the number of utterances of the respective participant to obtain a qualitative standard of a participant's mean utterance. Calculation of pq index is demonstrated in Table 4.q.

Table 4.q. Calculation of a pq index of an individual's performance in a tetralogue.

The number of utterances of participant P in tetralogue T: **up**					
Frequencies of utterances scored with:		**Frequencies of utterances scored with surplus combinations:**		**Frequencies of utterances scored with unpaired performances:**	
Idt:	I part.	Idt & Te:	IT part.	a single Te:	TS part.
Op:	O part.	Idt & Re:	IR part.	a single Ep:	ES part.
Te:	T part.	Op & Te:	OT part.	a single Re:	RS part.
Ep:	E part.	Op & Re:	OR part.	———	+
Re:	R part.	Te & Re:	TR part.	Subtotal 3:	**ST3**
———	+	Ep & Re:	ER part.		
Subtotal 1:	**ST1**	Idt & Te & Re:	ITR part.		
		Op & Te & Re:	OTR part.		
		———	+		
		Subtotal 2:	**ST2**		
Participant's philosophical quality in T: **pq** = (**ST1** + **ST2** – **ST3**) / **up**					

An example illustrates this formula:

Imagine participant J in a tetralogue videotaped in A. This tetralogue lasts 37 minutes and counts a total number of utterances of 307. J contributes with 79 utterances and scores as shown in Table 4.r.

Table 4.r. Calculation example of the pq index of J's performance in tetralogue A.

The number of utterances of participant J in tetralogue A: 79					
IDT:	14	IT:	4	TS:	14
Op:	17	IR:	5	ES:	3
Te:	31	OT:	8	RS:	7
Ep:	5	OR:	8	———	+
Re:	25	TR:	12	ST3:	24
———	+	ER:	-		
ST1:	92	ITR:	3		
		OTR:	4		
		———	+		
		ST2:	44		
J's philosophical quality in A: pq = (92 + 44 – 24) / 79 = 1.41					

Imagine that in the same tetralogue A, participant R performed a subtotal ST1 + ST2 + ST3 = 105. R performed this subtotal in 118 utterances. The pq of R is: 105 / 118 = 0.89.
In tetralogue A performed J a higher pq than R.

Using this formula, pq indices can be calculated for all participants within a tetralogue. The resulting numerical index reflects the philosophical quality of a participant's performance in an average utterance during a tetralogue.

In this project, 216 pq indices (of individuals) were calculated, ranging from 0 as the absolute minimum, to 3.94 as the observed maximum. Other statistics are: mean = 1.38; standard error of mean = 0.04; mode = 1.5; standard deviation = 0.71; variance = 0.50. Four participants scored 0, the lowest possible index. One of these participants made no utterances, another made utterances without any indicator assignation, and two participants made utterances assigned to Ep only. Theoretically, the maximum pq is 13. This can be obtained if a participant scores Idt, Op, Te, Ep, and Re in all his utterances [5 (maximum number of indicators) plus 8 (maximum number of valued combinations) in each utterance].

4.5 *Constructing an index for the philosophical quality of a tetralogue*

The philosophical quality of a tetralogue (PQ) has to be distinguished from the philosophical quality of individual contributions (pq). The PQ index of a tetralogue is more than the sum of individual pq's because the philosophical quality of a group performance should also take into account the combined forces of participants to arrive at a philosophically qualified outcome. The PQ index of a tetralogue will be based upon the pq indices of individual performances, but also must take into account: a) The combined efforts of the participants in the process of finding an answer on philosophical questions; and b) The dialogical features of the tetralogue as a whole.

The philosophical quality of a tetralogue is partly determinated by philosophical qualified contributions. But it must be allowed that the philosophical quality of a tetralogue is independent of the performance level of the lowest scoring participant, or participants who do not express themselves. Furthermore, the quality is conditioned upon the quantity of qualified utterances, i.e. those utterances of all participants that have received any indicator score. Potentially disturbing non-philosophical tetralogue characteristics may be: a) Duration, signifying the opportunity to 'score'; b) ratio between the number of utterances and duration; and c) A non-standard number of participants. These factors will be discussed later on.

A numerical PQ index for individual tetralogues should take into account the following factors:

1. Neglect of the lowest valued individual contribution
2. Philosophical qualified extent of a tetralogue

3. Combined forces of participants to arrive at a philosophical qualified outcome or the dialogical nature of a tetralogue
4. Number of participants that differs from the traditional set of four
5. Duration and speed of utterance production.

These factors are considered in detail below. The formula for the numerical index to be proposed will reflect the philosophical quality of a tetralogue.

Neglect of the lowest qualified contribution

A tetralogue is always composed of individual contributions. The construction of a PQ formula, therefore, has to take into account pq indices reflecting the individual philosophical performances of participants. However, the qualification of a tetralogue as a whole should depend on the best performing participants. The contributing pq indices are consequently restricted to those of the three highest scoring participants. Excluded from the quartet were students who did not participate sufficiently (less than 15 utterances) or had the lowest pq performance score. This restriction makes comparison possible with the few tetralogues that involved three or five participants. In cases of three participants, all contributions were taken into account, in cases of five participants, the three best contributions qualified.

The philosophically qualified content of a tetralogue

The philosophically qualified content of a tetralogue carries a quantity of qualified contributions. The quality becomes apparent in the pq indices of the three highest performing contributors of a tetralogue, reflecting participant's proportionate contribution in one utterance. To compute a PQ index of a whole tetralogue, the *mean* of these three pq indices was used. The quantity encompasses the total number of philosophically valued utterances: all utterances qualified with one or more indicators, and will be influenced by the mean quality of the three best performing participants. This quantitative expression takes into account all qualified utterances and to some extent the combined forces of participants to arrive at a philosophical qualified outcome. The PQ index concerns utterances of all participants, regardless of their relative pq level and number of participants, and includes single appearances of Te, Re, Ep, and of An: anecdotal quality. Indicators of anecdotal quality were excluded from the calculation of individual pq indices because they do not illustrate a philosophically qualified thinking pattern on the level of an *individual* and has its value only on a group level. Performances of anecdotal quality are believed to upholster and potentiate concepts and may subsequently direct or line up general thinking patterns of the tetralogue's community of inquiry. The number of utterances in which anecdotal qualities were expressed must be taken in account and be part to the frequencies of qualified utterances. As a result, one part of the PQ formula will consist of the product of the mean of the three highest pq indices and the number of philosophically qualified utterances in one tetralogue.

Dialogical events

A tetralogue carries in essence a collection of 'dialogues'. Individuals, through their contributions to the tetralogue, interact with each other to arrive at a philosophical qualified outcome. This outcome will be expressed partly by the number of jointly performed qualified utterances and partly by the paticipant's interactions. Interaction is mainly in the form of certain successions of utterances of several participants that constitute qualified combinations of indicator performances.

The first part of the PQ formula consists of the number of simple qualified utterances and does not take into account properties of indicator combinations performed by different participants over successive utterances. These properties or effects can be materialised by considering sequences in the scored philosophical indicators and result in a number of dialogical events. The sequential character of performances can be examined in the same way as the calculation of individual pq indices. As discussed earlier, some combinations are more important than others. Of interest in performances are sequences of [Te and/or Re] immediately following [Idt and/or Op], since these demonstrate intellectual inquiry in the group. If a sense for ambiguity, intellectual uncertainty, vagueness, relativity, openness, and wondering is followed by the autonomous production of tentative or speculative and reasoned thoughts, this could be indicative of the process assumed in the theoretical framework outline in Chapter 2. Combinations with Ep that were identified to expose a surplus value for individual pq indices were not valued in the dialogical events because Ep refers only to an individual performance. The same holds for most of the sequence combination Te-Re. This combination refers mainly to individual trials.

For each tetralogue, two time series were created: one indicating the presence of [Te and/or Re]; the other the presence of [Idt and/or Op]. By lagging one time series with one utterance in relation to the other time series, it was possible to calculate, per tetralogue, the number of dialogical events [Idt and/or Op] immediately followed by [Te and/or Re]. This qualified number of dialogue-events (D-events) forms the second part of the PQ formula for the whole tetralogue.

Now the components of the PQ index can be combined:

PQ = Mean pq index * Qualified utterances + Dialogue-events

Where:
- PQ refers to the philosophical quality of a tetralogue.
- Mean pq index refers to the mean of three highest pq indices only, based on the presence of five indicators [μ(pq1, pq2, pq3)].
- Qualified utterances refer to all utterances qualified with one or more of the six indicators.
- Dialogue-events refer to the number of times that [Idt/Op] is followed by [Te/Re] over two successive utterances.

Imagine two tetralogues: Tetralogue A and Tetralogue B (see Table 4.s).

Table 4.s. Calculation example of the PQ index in two tetralogues: A and B.

	Tetralogue A	**Tetralogue B**
Number of participants	4	4
Participant expressing less than 15 utterances	-	1
Pq indices of the three highest performing participants	pq1: 1.41 pq2: 0.90 pq3: 0.89	pq1: 3.35 pq2: 2.12 pq3: 0.97
Mean pq index μ(pq1, pq2, pq3)	1.07	2.15
Number of qualified utterances	199	130
Number of dialogical events	32	34
PQ index tetralogue	**1.07 * 199 + 32 = 244.27**	**2.15 * 130 + 34 = 313.07**

As a result, Tetralogue B shows a higher PQ index than Tetralogue A.

Empirically obtained values of PQ

Of the 70 tetralogues, the PQ index for two tetralogues could not be calculated as they were performed by two participants only. The PQ indices for the remaining 68 tetralogues were calculated. Their values varied from 51 to 479, showing a mean PQ of 266 and a standard deviation of 109. The higher the PQ index: the higher the philosophical quality of a tetralogue.

With respect to the constituting elements, the mean pq index of the three highest performing participants, the number of qualified utterances, and the number of dialogical events, the following figures are relevant: the maximum μ(pq1, pq2, pq3) appeared to be 3.32. However, this maximum comes from a tetralogue with low reliabilities counting a small number of utterances (86) (see Section 4.4.3) and produced the two highest individually performed pq indices. For the other 67, μ(pq1, pq2, pq3) varied from 0.56 to 2.68, a mean of 1.57, and standard deviation of 0.53. The number of qualified utterances varied from 58 to 281, with a mean of 155 and standard deviation of 55.93. The number of dialogical events varied from 4 to 48; mean: 27; standard deviation: 12.03.

Speed of utterance production

Apart from the quality and quantity of participant's contributions and their interactions, tetralogues differ in their duration and socio-cultural bias. Some tetralogues contained a relatively small number of utterances because students had the ability to listen to other participants without interrupting, allowing extremely long utterances containing one or more indicators. On the other hand, some tetralogues comprised an extremely large number of utterances per minute. Participants often interrupted each other. Sentences were short and excelled in over-simplified and blunt remarks that were sometimes ambiguous in a suspect way. As opposed to the 'long-turn' tetralogue, utterances in these 'in-a-jiffy' tetralogues

presented limited slight opportunities to score qualified expressions. As long as in one utterance each indicator can be scored once at the most, the ratios between number of utterances and duration offer different opportunities to score indicators. Some might argue that these effects influence the philosophical quality of the tetralogue because a large number of scarcely qualified utterances will load the first expression of the PQ formula to the same extent as a small number of rich, qualified utterances. The number of qualified utterances depends on duration ($r = 0.45^{**}$).

Duration

Tetralogues differ in their duration. Since long tetralogues consisted of more utterances and consequently provided more opportunities to score, a positive correlation was expected between PQ indices and tetralogue duration. Except for two very short tetralogues, the minimum duration was found to be about 25 minutes. And, since all tetralogues were executed in the natural setting of a teaching period unit, or within an agreed duration (Section 4.2.2), only four tetralogues exceeded 47 minutes: respectively, 48 minutes, 49 minutes, 50 minutes, and 53 minutes ($N = 70$). The mean duration is 38.8 minutes and the standard deviation is 7.6.

PQ turned out to be related to duration ($r = 0.34$, $p < 0.01$); this might be caused by the inclusion of some tetralogues with very short duration. 75% of the tetralogues had durations between 30 and 45 minutes. This duration can be considered as standard and is concordant with the effective duration of a lesson at secondary school. For the group of tetralogues with standard-durations the correlation between PQ and duration dropped to $r = 0.27$ ($p = 0.05$), marginally significant. Anyone estimating a tetralogue's quality, is not likely to look at their watches. This is supported by the non-significant correlation between estimated quality and duration. As will be explained in a next section, simple estimates of philosophical quality, performed by tetralogue groups, were made at an early stage in this study. It can be concluded that duration has some influence on PQ, but this influence can be contained to an acceptable level by restricting tetralogue duration to a 30 to 45 minutes range.

4.6 *Support for the validity of the PQ and pq indices*

Collective and individual thinking patterns expressed in a tetralogue have measurable philosophical qualities, which are supposedly picked up by PQ and pq indices. To check the plausibility of this hypothesis, PQ and pq indices will be empirically related to independently obtained measures of qualities expected to show concordance. If concordances are found this supports the validity of PQ and/or pq indices as measures of philosophical quality. Chapter 6 will contain a more elaborated validity analysis.

Calculated PQ indices will be compared with previously obtained global estimates of their philosophical quality. Indices of individual performances (pq) will be compared with pq indices of the same participant in different tetralogues.

Convergence between the PQ index and previous PQ estimates

The PQ formula produces a quantitative index of the philosophical quality for each completed tetralogue. To be valid, this index should be corroborated by other estimates of philosophical quality. At an early stage of this project all videotaped tetralogues were judged on a [0-5] point-scale by the author and a representative sample of ten of them also by a second experienced philosophical judge. Because co-judging was executed in cooperation with the author to establish the appropriate criteria, no inter-rater agreement between independent judges was determined. Fortunately, inter-rater consensus was easily reached. Judging was an overall estimation executed on the basis of watching the video only. During that judgement of a tetralogue, criteria for philosophical quality were taken into consideration based on five parameters derived from the definition and the main features of philosophy: 1) The production of autonomous thoughts by participants; 2) The intellectual distance to the philosophical question; 3) The ability to generate new questions, wonderings and uncertainties; 4) The understanding of the complexity of the matter or growing consciousness of the degree of complexity; and 5) The ability to persist in systematic searching (see also Chapter 2 and Section 4.4.1). The judged presence or absence of each of these five criteria generated a score on a [0-5] point scale and indicated the *estimated philosophical quality*.

Convergent validity between the calculated PQ index and the early estimate of philosophical quality was anticipated. On the other hand, the correlation between the two sets of measures should not be too high, because the PQ index is the result of objective, detailed, theory-based procedures and is expected to out-perform any crude judgment. Moreover, the early estimates of tetralogue quality were based on overview criteria not included in the calculation of PQ indices. The Pearson correlation between the calculated and the estimated philosophical quality of 68 tetralogues was: $r = 0.54^{**}$ ($p < 0.01$; one-tailed test). This correlation proves that the PQ index, as derived by the methods described in this chapter, converges with expert estimates. Moreover, this correlation is low enough to show distinctiveness between the two measures. The size of this validity coefficient might be spurious as both measures might be influenced by duration of the tetralogue. However, the partial correlation between PQ and the estimated tetralogue quality hardly dropped (from $r = 0.54$ to $r = 0.50$) when duration is partialled out.

Individual's pq index

Tetralogue behaviour is a function of individual's characteristics and of the discussion situation. However, when pq is a good measure of a characteristic that the individual brings to the situation, considerable trans-situational consistency in the tetralogue behaviour should be observed. Calculated pq indices are supposed to represent philosophical qualities characteristic of individual's performances in tetralogues. If the pq index represents any constant aspect of an individual, this should be reflected in some degree of convergence between the pq indices in two different tetralogues. As mentioned in the last section, 27 youngsters participated in two (or more) tetralogues; two of them in non-scorable tetralogues. Comparison

of the pq indices of the 25 contributors over two tetralogues results in a correlation coefficient of $r = 0.61^{**}$ ($p < 0.01$, one-tailed). This significant and high correlation demonstrates considerable convergence of individual performances over two tetralogues and supports the hypothesis that pq indices are characteristic of individual performances. Almost all participants ($N = 21$) acted in two different tetralogues with different themes but with exactly the same participant composition. The correlation coefficient for the pq indices of these 21 participants over the two tetralogues is: $r = 0.64^{**}$ (one-tailed). This suggests the instrument used to assess individual level of philosophical quality is measuring something similar in two different discussions with similar participant composition.

If this observed trans-situational convergence were the result of some constant aspect of the tetralogues, it should also be observed between different co-participants contributing to the same pair of tetralogues. A correlation was calculated between pq indices of participants in the first tetralogue with pq indices of a random selection of co-participants in the second tetralogue. No significant correlation could be demonstrated. Thisprovides strong evidence that the pq index reflects an individual characteristic.

5 Conclusions

In this chapter, the philosophical quality of thinking patterns was explored using the tetralogues as a standardised instrument for elicitation and investigation. Six indicators derived from a theoretical framework identify philosophical qualities of expressed thinking patterns. Quantitative measures of philosophical qualities of individual participants (pq index) and of the tetralogues (PQ index) were established. Objectivity of scoring, manifestations of reliability and validity were investigated. The following conclusions can be drawn from this research:

1. Philosophical quality of thinking patterns can be judged on the basis of six indicators: Indecisive thinking (Idt), Openness (Op), Tentative behaviour (Te), Epistemical position (Ep), Reasoning quality (Re), and Anecdotal quality (An).

2. Indicator presence in the utterances expressed in a tetralogue was found to be objectively scorable by judges provided with a philosophical background and training.

3. All indicators proved insensitive to random fluctuations and so formed a reliable basis for assessing philosophical qualities of individual and jointly performed thinking patterns performed in tetralogues .

4. Reliabilities of tetralogues were not related to philosophical theme discussed, participant's age, educational level or life experience.

5. The five indicators used to judge individual performance converge; their internal consistency was computed: $\alpha = 0.80$.

6. Philosophical qualities can be expressed through numerical indices for individual participants in tetralogues (pq) and for entire tetralogues (PQ).

7. PQ measures, as derived from the methods described in this chapter, correlate substantially but distinctively with expert estimates made independently.

8. The pq index measures of individual participant's contributions in two different tetralogues show trans-situational convergence.

9. The results show that pq and PQ indices meet the prerequisites for being valid measures representing the philosophical quality of an individual performance (pq) and group performance (PQ).

5 Similarities and differences between philosophical themes

The tetralogues in this study have been conducted in a consistent, standardised manner. They begin with a philosophical question and are performed by a group of (mostly) four participants and a chairperson. Differences between tetralogues are potentially caused by: a) type of initial questions and elaborated themes; b) type of participants and group composition; and c) type of discussion and the chairperson. This chapter adresses similarities and differences between philosophical themes. Similarities and differences between participants will be explored in Chapter 6. In Chapter 7, variations related to the type of discusion and the chairperson are reviewed.

Tetralogues were designed as instruments to measure philosophical quality of discussions. This requires that participants respond to a standardised situation. Only then can differences in responses between participants be attributed to something that the participant brings to this situation (e.g. his or her philosophical talent). Standardisation of the test situation of tetralogues has been described in Section 4.2 in the form of rules and procedures. The purpose of these rules and procedures is to keep all irrelevant, but potentially influential factors constant. The philosophical theme or topic of tetralogue discussion has not been kept constant. So, it is of interest to check whether the type of philosophical theme has not biased the philosophising performance.

This chapter focuses on differences between themes in elicited performances. Every tetralogue is an elaboration of an initial, key question by participants. Derived from this question, several philosophical themes and notions were explored. The common denominator of a tetralogue's initial question is: 1) the non-existence of a definitive answer (i.e., an open question); 2) the potential of a systematic inquiry; 3) its rooting in concrete occurrences; and 4) its voluntary realisation by participants. Differences between themes were introduced in Chapter 3. These differences will be discussed in more detail in Section 5.1. Any analogy between tetralogue themes and themes in academic philosophy is explored in Section 5.2. In Section 5.3, similarity in the character of inquiry of the thematically different tetralogues will be examined (i.e., a tetralogue's answer-finding procedure). Analogies between tetralogues with different philosophical themes will be evaluated with respect to their philosophical qualities. Finally, in Section 5.4, variations in tetralogue characteristics will be identified and their function in philosophical topics reviewed.

1 Different themes of the tetralogues

Each tetralogue begins with an initial key question raised and formulated by one participant and successively agreed upon by all participants. This question should reflect a philosophical theme that is embedded in a concrete experience or anecdote. Sometimes questions are ambiguous and generate two or more philosophical themes. In this study, tetralogues cover four main themes: 1) metaphysics and

epistemology; 2) anthropology; 3) ethics; and 4) meaning and demarcation problems. If a tetralogue addressed more than one theme, it was assigned to the dominant theme. Two of the 95 tetralogues addressed multiple themes to equal extent and were placed in a 'rest' category.

Distribution of tetralogue themes may be arbitrary for three reasons: 1) observable discriminative criteria had yet to be detected; 2) thematic questions are always ambiguous and may generate two or more themes and philosophical notions in addition to those covered by the initial question; 3) the historical development of academic philosophy has created a variety of different thematic specialities and different classifications of philosophical themes. Themes from academic philosophy will be discussed in Section 5.2.

Themes embedded in questions and elaborated upon in tetralogues

As demonstrated in Chapter 3, philosophical themes reflected by the initial question can be distinguished from other themes and notions elaborated upon in tetralogues. The proportion of themes generated by the initial question and those elaborated subsequently in the philosophical discussion can be illustrated by referring to *Jong & Wijs* (Rondhuis, 2001). In this booklet, some 50 philosophical discussions are described and their elaborated themes are categorised in a register. There, initial questions can be assigned to the four mentioned themes: 13 (26%) epistemological or metaphysical questions; 18 (36%) anthropological questions; 6 (12%) ethical questions; and 13 (26%) meaning questions or with reference to demarcation problems.

In all philosophical discussions, more than one theme was elaborated upon. Consequently, these 50 discussions generate more than 50 themes: 37 (74%) of which were assigned to epistemology or metaphysics; 28 (56%) to anthropology; 13 (26%) ethics; and 19 (38%) to meaning and demarcation problems. The assignment of themes to tetralogues was based on the predominant philosophical theme elaborated. The high percentage of epistemological or metaphysical themes is most remarkable. However, since key questions are not unambiguous, the theme elaborated upon is subsequently unpredictable. In some cases, unrelated themes were developed in response to key questions. In these instances, tetralogues were assigned to the theme elaborated upon in the tetralogue and not to that of the initial question. Therefore, in this study, a distinction cannot be made between themes of initial key questions and of those elaborated upon in a tetralogue.

Metaphysics and epistemology

Metaphysical questions are concerned mostly with fundamental questions about the foundation, structure and interpretation of reality. Epistemological questions examine knowledge and its relationship with reality. Many initial key questions have a metaphysical dimension that is thought through, even if such dimension cannot be initially detected. Traditionally, metaphysical and epistemological themes

contain a high degree of abstractness and seem to be rather sophisticated. However, in tetralogues, such themes are rooted in the events of daily life. They occur frequently. They refer to the following subjects and oppositions: truth and perception; identity and authenticity; existence and creation; knowledge and beliefs; time, space and eternity. In this study, examples of initial key questions are:

1. *Do colours exist if no one can see them?*
2. *Does St. Nicolas (Sinterklaas) really exist?*
3. *What is the difference between experiencing an event and reading about it?*
4. *Does America really exist?*
5. *Can reality be exchanged for virtual reality or for a dream?*
6. *What is the identity of a renovated piece of art?*
7. *What kind of reality represents a dream?*
8. *Do paranormal phenomena really exist?*
9. *What is meant by chance?*
10. *Does the universe end?*
11. *Do deep-frozen people live forever?*

Questions about dreaming appear to be very popular amongst 12 and 14 year old children. However, since dreams are unverifiable, these tetralogues have a large number of tentative remarks and a high frequency of cognitive verbs, and first person, singular remarks (see Section 5.4). Some tetralogues focusing on dreaming develop in an anthropological sense and were assigned to the following category of themes.

Anthropology

Anthropological questions comprise all those about human existence, including mind-body problems, differences between man and animal, affections and emotions, life and death, free will, and individual identity. Initial questions that are frequently encountered in this study are:

1. *Am I still myself if my body (or name) changes?*
2. *Is it possible to be two individuals in one person?*
3. *Is a deep-frozen person alive?*
4. *Can a pet (statue or ball) live?*
5. *What is meant by 'life' if you are not able to perform essential functions like communicating?*
6. *Is love physical or mental?*
7. *Is it possible to fall in love with an alien or robot?*
8. *Can a computer become world champion chess playing?*
9. *Can a painting monkey be considered an artist?*
10. *Given the choice between two different paths, in what direction would a two-headed snake choose to go?*

Some tetralogues begin with questions on subjects of aesthetic origin, such as 'Can animals produce art?' Although these questions may require some normative

judging, children and youngster in this study tend to develop anthropological themes in response to these questions. Apparently, participants were not familiar with artistic quintessence, or simply preferred to explore the anthropological dimension of the initial question.

Ethics

In this study, ethical themes are developed based on questions about decisions made by individuals or society on issues of good and bad. Key questions may include:

1. *Can a senile grandparent still make decisions?*
2. *If possible, is the exchange of the human body parts between different persons permissible?*
3. *May abortion and euthanasia be applied in special circumstances?*
4. *Can money make people happy?*
5. *How should personal interests be balanced with respect to those in a group?*
6. *Why should you have good manners?*
7. *What is a sin?*
8. *What is the sense of meaningless violence?*

Many questions have clear meanings and are assigned to this category because of their social significance and public connotations. According to many statements in Dutch journals and by the current (= 2005) Dutch Prime Minister, ethical questions are important during adolescence in a well-developed and integrated society. Schools may table ethical questions on the agenda when pursuing educational goals in philosophy lessons. These goals are certainly not challenged in the tetralogues. When dealing with ethical themes three characteristics must be emphasised because of their social influence: 1) the prescriptive and (often) commonplace character of many ethical themes; 2) the transition from experiencing to thinking and *vice versa*; and 3) the relationship between thinking and acting.

Prescriptive and commonplace character

All sorts of do's and don'ts dominate the morality domain of youngsters. As a consequence, they may loose their authentic thinking patterns when experiencing abundant prescriptions and presuppositions. For example, because one is supposed to act against vandalism and to sympathise humankind and nature, thinking patterns should be developed accordingly. Discussions on ethical topics are generally dominated by clichés and overworked phrases. Political correctness is often an important factor driving these discussions, even if no ethical key question would be involved. For example, in a group of 11 to12 year-old children philosophising about the demarcation between small and large, participants explored the concept of size by thinking about small and large people. They often argued in ethical terms, for example: 'I think it is not good to condemn people for their size'. The linguistic use of many words tends to be ambiguous but with a clear ethical dimension. *Small* refers not only to a physical size but also to a mental attitude

that is considered negative. Political correctness and overworked confrontation with issues of good and bad often hide clear arguments and avoid lucid reasoning. Therefore, the risk of running into politically correct answers on initial key questions was checked at the beginning of each tetralogue. If positive, the initial question was replaced by another question: sometimes, another ethical question, or a question from another category.

Transition from experiencing to thinking

A smooth transition from experiencing to thinking, and vice versa, provides the grounds of a specialised philosophical movement called 'pragmatism'. According to Dewey (1984), children make slight distinctions between thinking and experiencing. Their smooth transfer from experience into thought and *vice versa* may refer to an inequality between comprehension and production (Berk, 1997). Young children are supposed to be unable to justify what they understand. Consequently, comprehension is like experiencing without production of verbal behaviour, whereas producing words refers to conscious thinking. Besides, children frequently wield a severe consequentiality and rigid logic with their thinking patterns (see also Chapter 3): what applies in the domain of experience consequently applies to the domain of thinking. Attribution of souls to things, empathy and role-taking seem to follow naturally from this line of cognitive behaviour (thinking and reasoning). For many children, a stellar object really lives and a rabbit can still *be* himself. In Chapter 3, Umi's role-taking attitude (Valuing poems) and her imagination of being a flea more than exemplifies this phenomenon. The non-intellectual capacity of empathy and the smooth transition between experience and thought is a component of philosophically qualified thinking patterns. It is supposed to be expressed in uncertainties, tentative behaviour, judging weird, wonderings, modalities and unusual word associations (see Chapter 3), and is consequently covered by the respective indicators.

Thinking and acting

Good thinking is not the same as good acting. A poor transfer from thinking to acting may cause undesirable disharmony between the two. Ethical questions attempt to cover this relationship. Judging the philosophical quality of thinking patterns evoked by ethical questions also explores the recognition of moral duty. Children's recognition and acceptance of moral duty does not automatically coincide their capacity for moral reasoning and justification (Matthews, 1994). They recognise more moral duties than they can justify. This complies with the fact that children understand more than they can communicate, or produce what is understood (Berk, 1997, p. 354). Recognition of moral duty is an experience preceding its justification. While identifying moral reasoning can assess justification, recognition of duty cannot. Ethical thinking patterns comprise more than cognitive reasoning and the offering legitimacy or justification. Empathy, recognition and acceptance of moral duty also appear to be at stake. Returning to examples of philosophical discussions arising from ethical questions like 'voting incorrectly' (see Chapter 3),

recognised expressions include the non-intellectual capacity of empathy like: children's wondering; their judgement of weird; sensing ambiguities, vagueness and uncertainty; reasoning through analogy; and their searching for logic where it does not appear. Such behaviour belongs to philosophical qualities of thinking patterns on ethical themes. As with other themes, they are evaluated by scoring philosophical quality (pq) indicators.

Meaning and demarcation problems

Questions of this category refer to the mysterious or cryptic relationship between words and things, and to demarcation lines between corresponding concepts. How is communication about things and events possible? In tetralogues, concepts are explored while searching for boundaries, connotations, associations and oppositions. Examples of this form of initial key question are:

1. *Do all things have their anti-poles?*
2. *What really is cold (large, beautiful ...)?*
3. *What exactly does scoring an unsatisfactory mark mean?*
4. *What is the difference between animal and vermin (in Dutch: non-animal)?*
5. *Can a prisoner feel free?*
6. *Who is, or can be, normal?*

This category includes also themes generated but not necessary implied by initial key questions.

2 Academic philosophy and tetralogue philosophy

Traditional distribution of philosophical themes

Assigning themes to tetralogues is not unique as the same process is common practice in academic philosophy. Traditionally, philosophy is subdivided into metaphysics, epistemology, anthropology, ethics, aesthetics, philosophy of language and science, social and political philosophy, and the philosophy of history, culture and nature. These subdivisions concern questions about reality, knowledge, the human being, good, beauty and truth, and problems of meaning and demarcation. This conforms with the classical classification styles of Kant, referring to truth, good and beauty, or more specifically: metaphysics, critique of pure reason, critique of practical reason, and critique of judgement. The bottom line is that all assignations of themes are arbitrary because of the lack of observable discriminative criteria and due to chance historical developments. Such developments can be exemplified by referring to short periods of time. In the period between 1985 and 2000, many structural changes took place at the faculty of philosophy in the University of Amsterdam: for example, the distribution of chairs in academic philosophy changed five times. Students had to adapt to changes in the classification or labelling of philosophical disciplines, even though available courses and

corresponding texts did not change. Despite the administrative changes, faculty and students were supposed to maintain traditional philosophical thinking patterns.

Differences between academic and tetralogue philosophy

Although tetralogues are manifestations of philosophising, they are not identical to academic philosophy. Academic philosophy comprehends reflection on its own history and often offers conceptual and theoretical structures to map reality. Philosophising youngsters do not. Rather, they search for the why and how, but stay away from the historical context, well-known philosophers and their specific range of ideas; nor do they attempt to constitute universal theoretical structures to map reality. Academic philosophy presupposes historical knowledge that is indispensable for understanding significant philosophical ideas and notions.
A third point of difference between academic philosophy and tetralogue philosophy is the apparent absence of certain types of philosophical themes in the latter, especially with respect to aesthetics and to the foundation of social sciences.
A fourth dissimilarity concerns the tangibility of themes, the smooth transfer from abstract into concrete levels of themes, and the spontaneity of thinking patterns. These aspects occur in tetralogues more than in regular academic philosophy.

Tangibility

Thinking patterns in tetralogues are evoked by and incorporate events of real life. They are rooted in tangible problems. When problems are offered in an abstract mode, youngsters easily translate them into concrete situations, magnified with connotations and association. Nelson's regressive abstraction can be subsequently executed. Thinking patterns of children and youngsters are not concerned with historical perspectives or by the consequences of well-formulated (and founded) concepts of reality. In contrast to academic philosophy, children keep less intellectual distance from philosophical themes. Intellectual distance leads thinking patterns into more abstract reasoning, while entanglement in real life and concrete events generally ends with spontaneous reactions, and leads to production of autonomous lines of thought. Some academic philosophers emphasise the significance of this character of concreteness, of some back-to-earth stance, and of thought experiments based on tangible problems that are designed to arrive at philosophically qualified thinking patterns (e.g., Hoffstadter, 1985; Nagel, 1974, 1990; Dennett, 1993; Wittgenstein, 1973).

Irregular experience

The significance of real life experience connected with ethical themes and based on a supposed relationship between experiencing and pondering, generates one of the hypotheses for this study. Irregular or special experience, like being adopted, handicapped or in prison, may drive individuals into pondering and producing thinking patterns about identity, free will, abortion, the value of perception, and

the value of the physical body. Differences between regular and irregular experienced youngsters will be discussed more in detail in Chapter 6.

3 Similarities in the character of inquiry

It is the aim of a tetralogue to identify qualified answer-finding procedures that are realised by participants in a collective inquiry, rather than qualifying its contents. Expressions of this inquiry are scored on their philosophical quality (PQ). Chapter 4 demonstrated that the reliability of these scores is independent of distinguished themes or topics. Although initial key questions and elaborated tetralogue themes may differ in their content, no differences are supposed to occur in the answer finding procedures. Tetralogues are designed to register and measure philosophically qualified thinking patterns through these answer-finding procedures by a community of inquiry *regardless* of the content of the initial question. Boundary conditions for initial key questions were: openness, intelligibility, concreteness, and been selected by tetralogue's participants as requiring inquiry. If different themes evoke thinking patterns with different philosophical qualities, different PQ indices will appear in processing tetralogues with distinguished themes.

In this study, 95 tetralogues were categorised according to their theme: 35 with 11 to 13 year-old children, 35 with 14 to 16 year-olds, and 25 with 17 to 19 year-old youngsters. From these: 34 tetralogues focus on epistemological or metaphysical themes; 40 tetralogues elaborate on anthropological themes; 12 tetralogues develop ethical themes; 7 tetralogues are based on questions of meaning and demarcation; while 2 tetralogues are assigned to a remainder category. All youngsters but the group of 17 to 19 year-olds qualify for investigation of philosophical quality. Relationships between type of initial questions and philosophical quality were based on calculations from 70 tetralogues of 11 to 16 year-olds. Similarities between theme groups in tetralogues with regard to philosophical qualities are shown in Table 5.a.

Although tetralogues were assigned to different themes, these proportions are very similar. The mean PQ indices of the four philosophical topics do not differ significantly: $F(3,64) = 0.17$, n.s. Nor do the number of qualified utterances differ over the thematically grouped tetralogues: $F(3,66) = 1.36$, n.s. However, the number of zero utterances, utterances without any indicator score, do differ significantly: $F(3,66) = 4.81$ ($p < 0.01$). Reason for this deviation is that the group of tetralogues focusing on anthropological questions generate a relatively large number of zero utterances. The characteristics of anthropological driven tetralogues will be described below.

It could be possible that the *philosophical quality* indicators - the building stones of PQ indices - differ between distinguished theme groups of tetralogues. There are no differences in the mean frequencies of separate indicator occurrences in thematically grouped tetralogues. Similarities between tetralogue theme groups with regard to the six indicators, constituting philosophical quality, are presented in Table 5.b.

Table 5.a. PQ indices and proportion of qualified utterances over tetralogues with four types of themes.

Themes	N	PQ index tetralogue			Qualified Utterances			'Zero' Utterances		
		Min.	Max.	Mean (s)	Min.	Max.	Mean (s)	Min.	Max.	Mean (s)
Total	70	51	479	266 (109)	58	281	154 (56)	4	137	52 (30)
a. Epistemology Metaphysics	26	51	479	270 (132)	60	238	153 (55)	15	98	47 (23)
b. Anthropology	32	90	473	261 (93)	58	281	164 (58)	4	137	63 (34)
c. Ethics	8	129	459	288 (110)	79	207	139 (51)	16	58	31 (14)
d. Meaning & Demarcation	4	171	290	243 (63)	77	156	110 (34)	7	33	25 (12)

Legend: Qualified Utterances: Number of utterances, qualified with one or more indicators
'Zero' Utterances: Number of utterances, not qualified with one or more indicators
(s): Standard deviation

Table 5.b. Distribution of mean frequencies of indicators over tetralogues with different themes.

Mean indicator frequencies (s)	Idt	Op	Te	Ep	Re	An	N
Overall Mean	43 (18)	28 (17)	80 (28)	58 (25)	48 (23)	5 (5)	70
a. Epistemology Metaphysics	49 (22)	27 (18)	78 (26)	57 (29)	46 (26)	6 (6)	26
b. Anthropology	41 (14)	30 (17)	84 (30)	58 (24)	50 (24)	4 (3)	32
c. Ethics	35 (15)	28 (13)	73 (29)	61 (20)	49 (20)	4 (4)	8
d. Meaning & Demarcation	36 (9)	19 (13)	65 (16)	54 (23)	39 (13)	7 (3)	4

Legend: Idt: Indicator of indecisive thinking
Op: Indicator of openness
Te: Indicator of tentative behaviour
Ep: Indicator of epistemological position
Re: Indicator of reasoning quality
An: Indicator of anecdotal quality
(s): Standard deviation

The proportion of mean indicator frequencies (Idt, Op, Te, Ep, Re and An) of a thematically grouped tetralogue looks similar to the mean frequencies for each indicator of all 70 tetralogues. Statistical analysis (ANOVA) confirms this. Mean frequencies of Idt, Op, Te, Ep, Re and An, when assigned to the four philosophical topics, do not differ significantly: F (3,72) values are, respectively: 2.04 (n.s.); 0.56 (n.s.); 1.00 (n.s.); 0.09 (n.s.); 0.49 (n.s.); 1.71 (n.s.). When utterances are scored, differences in tetralogue content apparently do not result in different PQ indices, number of philosophically qualified utterances, or types of indicator combinations. There is a common pattern of relative indicator frequencies present in all tetralogues irrespective of the initial question and content, which pattern may be considered as characteristic for all tetralogues. These similarities lead to the conclusion that a tetralogue's general philosophical character is independent of its thematic content. This character can be identified as tetralogue's philosophical inquiry.

4 Variations in tetralogue external characteristics in function of philosophical themes

Elaborated themes in different tetralogues

Apart from analogies in discussions with different philosophical themes, variations in tetralogue characteristics can be identified in the function of their philosophical topics. Philosophical topics can be investigated by comparing formal properties of tetralogues, such as duration time, number of utterances, and number of interventions by the chairperson.

Table 5.c. Formal properties of thematically grouped tetralogues.

	N	Mean Duration (s)	Mean number of Utterances (s)	Mean number of Interventions (s)
Total	70	38 (8)	206 (79)	37 (17)
Epistemology Metaphysics	26	37 (9)	201 (69)	42 (19)
Anthropology	32	39 (7)	227 (88)	37 (16)
Ethics	8	36 (4)	171 (63)	25 (17)
Meaning & Demarcation	4	35 (9)	135 (38)	32 (7)

Legend: (s): Standard deviation

No relationship could be detected between a tetralogue's philosophical topic and its duration: $F(3.67) = 0.80$ (n.s); nor do the number of utterances and number of chair interventions differ significantly between the four topics or themes: respectively, $F(3,66) = 2.65$ (n.s.) and $F(3,66) = 2.38$ (n.s.) (see Table 5.c).

Preferences of participants

One of the conditions for a tetralogue is the free choice of theme. To facilitate participants' choice a topic list is presented, but participants also proposed topics of their own. The collectively choosen topic can therefore be considered as a preference of the tetralogue group. It is of interest to learn whether preferences are related to background characteristics of participants: age, level of education, life course and gender dominance in the group. The groups' preferences in relation to their background variables are shown in Tables 5.d, 5.e, 5.f and 5.g.

Popularity of themes

Despite the small number of tetralogues covering ethical topics or meaning and demarcation problems, some inter-topic differences may be noticed. At first glance, a relationship appears between age groups and themes: metaphysical and epistemological themes seems popular among 11 to13 year-olds, anthropological themes amongst 14 to16 year-olds, while ethical themes or those concerning meaning and problems of demarcation seem to be relatively popular among 17

Table 5.d. Preferences for philosophical themes in different age groups.

	11-13 year olds		14-16 year olds		17-19 year olds		11-19 year olds	
	N	% valid	N	% valid	N	% valid	N	% valid
Epistemology Metaphysics	16	46	10	29	8	32	34	36
Anthropology	13	37	19	54	8	32	40	42
Ethics	4	11	4	11	4	16	12	13
Meaning & Demarcation	2	6	2	6	3	12	7	7
Rest category	-	-	-	-	2	8	2	2
Total	**35**	**100**	**35**	**100**	**25**	**100**	**95**	**100**

Table 5.e. Preferences for philosophical themes in groups of different educational level.

	HAVO-VWO		VMBO		All tetralogues	
	N	% valid	N	% valid	N	% valid
Epistemology Metaphysics	19	40	15	31	34	36
Anthropology	14	30	26	54	40	42
Ethics	10	21	2	4	12	13
Meaning & Demarcation	4	9	3	6	7	7
Rest category	-	-	2	4	2	2
Total	**47**	**100**	**48**	**100**	**95**	**100**

Table 5.f. Preferences for philosophical themes in groups with different life course.

	Regular life course		Multiform life course		All tetralogues	
	N	% valid	N	% valid	N	% valid
Epistemology Metaphysics	21	43	13	28	34	36
Anthropology	19	39	21	46	40	42
Ethics	7	14	5	11	12	13
Meaning & Demarcation	2	4	5	11	7	7
Rest category	-	-	2	4	2	2
Total	**49**	**100**	**46**	**100**	**95**	**100**

Table 5.g. Preferences for philosophical themes in groups different in gender dominance.

	Tetralogues predominated by males		Tetralogues predominated by females		Tetralogues with equal numbers of males and females		All tetraloguess	
	N	% valid	N	% valid	N	% valid	N	% valid
Epistemology Metaphysics	18	37	6	38	10	37	34	36
Anthropology	20	41	5	31	14	52	40	42
Ethics	3	6	5	31	3	11	12	13
Meaning & Demarcation	6	12	-	-	-	-	7	7
Rest category	2	4	-	-	-	-	2	2
Total	**35**	**100**	**35**	**100**	**25**	**100**	**95**	**100**

to 18 year-olds. However, the observed differences in topic choice between age groups are indistinguishable from random fluctuations. Multivariate analysis on the four dependent variables (age, educational level, life course and gender composition in tetralogue groups) shows an overall difference between four topics (Wilk's Lambda is significant, $p < 0.05$). Follow-up univariate F-tests show that the four topics differ with respect to level of education and gender composition of the tetralogue group. Ethical themes are popular within female-dominated groups and among higher educated participants. Anthropological themes seem to be favoured by lower educated participants and in groups with an equal number of girls and boys. Surprisingly, when comparing youngsters with irregular life course to regular youngsters, the proportion of tetralogues focusing on ethical themes with moral implications for human physical and mental existence is similar in this study (13.7%, N: 45 and 13.3%, N: 51, respectively).

5 Conclusion

To compare individuals in situations in which philosophical quality is expressed, tetralogues are standardised although they may differ in features like: duration time, number of utterances, number of interventions, and topics. Tetralogues are initiated by initial key questions voluntarily selected by their participants. Key questions can be assigned to the following categories: 1) metaphysics and epistemology; 2) anthropology; 3) ethics; and 4) philosophy of meaning and demarcation problems. This classification is broadly in accordance with philosophical traditions. Questions arise as to the relationship between a tetralogue's content and the philosophical quality it elicits.

Different topics or themes do not lead to significantly different philosophical qualities, nor to differential indicator patterns. The different themes are also not associated with differences in duration, number of utterances or number of chair interventions.

Since philosophical qualifications of tetralogues do not vary widely with regard to the four topics, and all thematically grouped tetralogues show a similar pattern of indicator frequencies, the general character of inquiry in an individual tetralogue may be established. Evidently, all four philosophical themes can be explored through similar answer-finding procedures and philosophically qualified thinking patterns.

Differences in preferences for key questions and themes can be noticed amongst participants in potential relation to their characteristics. Age and type of life course show no relation. Girls and youngsters with a low level of education favour anthropological initial key questions. Ethical themes score relatively high in female dominated groups and among higher educated participants. A convenient set of topics may be established with respect to different categories of participants to evoke philosophically qualified answer-finding procedures and thinking patterns.

6 Participant's characteristics in relation to philosophical quality

Similarities and differences between tetralogues regarding their philosophical quality can be identified and interpreted from different perspectives. In chapter 4 and 5 many analyses focused on the philosophical quality of group performances. Here, the focus will be on the philosophical quality of individual performances in relation to other characteristics of tetralogue participants. On theoretical grounds, a relationship can be expected between certain participant characteristics and their performed philosophical quality. These expectations will be used to further corroborate the validity of the pq index, in ways to be described below. Prerequisites for validity of the pq index were already shown to be fulfilled in chapters 4 and 5. Tetralogues could be scored objectively, and the indicators on which the pq index is based are demonstrated to be reliable and insensitive to the theme of discussion. Moreover, the pq index was shown to be a trans-situationally constant and highly individual characteristic.

In this chapter, the attempt of validation will be deepened and extended by: 1) a discussion on the concept of validity (6.1); and 2) a search for replication of the theoretically based relationships in empirically collected philosophically qualified thinking patterns. The nomological network around philosophical quality that forms the foundation for hypotheses to be tested is given in Section 6.2. To investigate correspondence between theory and empirical findings, participants were categorised according to selected relevant characteristics. Recruitment of participants and groups, measuring participant's characteristics, and a plan of analyses to examine possible relationships between characteristics and pq and PQ indices are presented in Section 6.3. Data and indices are processed statistically. Results of analysed relationships between characteristics and pq and PQ indices will be presented in Section 6.4 with respect to individual and group performance. Apart from the anticipated relationships between philosophical quality and selected characteristics, other relationships may also emerge concerning indices for philosophical quality. Such potential relationships will be explored in Section 6.5. Finally, a number of issues raised by the results will be discussed in Section 6.6.

1 Validation of the tetralogue as a measure of philosophical quality of thinking patterns of individual performances and of group performances

Construct validity and construct validation

The concept of construct validity expresses that an attribute, designated by a theoretical term, exists; and that measurement of this attribute can be performed with a given test because the test scores are causally affected by variation in the attribute (Borsboom et al., 2004). This statement is in line with classics concerning construct validation (Crohnbach & Meehl, 1955, Campbell & Fiske, 1959, Messick,

1989, Standards for Educational and Psychological Tests, 1999). So, a chain of events has to be established, beginning with *Event 1*, the existence of a concept in reality, through *Event 2*, theoretical terms and derived relationships; and ending with *Event 3*, test scores. Validation establishes whether this chain of events is solid. Validation is not a purely methodological enterprise: it is more than finding out the meaning of the measurement, looking for replication of the nomological network in test scores. To rely only on the epistemological process of meaning is to neglect the ontological claim referring to the relationship between theory and practice of measurement. Somewhere, the theoretical concept of philosophical quality has to attach to reality and subsequently, causal effects of this concept have to be conveyed on test scores. Only the truth of the ontological claim guarantees the epistemological access that will make the process of validation possible. Validity based on an epistemological criterion (meaning) is a match between empirical and theoretical relationships. If theoretical and empirical relationships do match, the correspondence corroborates the theory. Other views exist. The match no longer constitutes validity since the reintroduction of realist metaphysics forces one to return to reference as the primary defining feature of validity. If a term is treated as referential but has no referent, then one is reifying terms that have no other function than that of providing a descriptive summary of a set of distinct attributes and processes (Borsboom *et al.*, 2004, p. 1065). The authors state that reliability of a test presupposes validity in ontological sense, contrary to ordinary psychometric procedures.

Constitutive claims for tetralogue's validity

A tetralogue's validity is based on two constitutive claims and on procedures of looking for replication and correlation in the tetralogue as measurement to corroborate a corresponding theory. A tetralogue's validity constitutes: 1) the existence of philosophical quality as construct, and 2) its causal impact on the tetralogue scores. Hypothesis 1 is that philosophical quality, as a concept or theoretical term, refers to something in the world; to a ground for differences in philosophical performance repeatedly observed over a 25 year period while philosophising with children and youngsters (Rondhuis, 1994, 2000). It is assumed that this experience is grounded in an existing attribute, although this stance may sound naïve from a realist metaphysical perspective. However, not much more can be said without becoming entangled in a web of metaphysical approaches that lie outside the scope of this study. The second hypothesis concerns relationships of causation. Correlations may indicate causality in an epistemological sense. However, causality is not a transitive relationship free of obligations, but runs from the attribute to the measurement. In order to show that tetralogue scores are valid measures of philosophical quality, there should be at least a hypothesis concerning the causal processes that lie between the variations in philosophical quality and the differences in test (tetralogue) scores. Causal relationships between the attribute of philosophical quality and measured indices for philosophical quality (pq and PQ) stress the foundation of philosophical quality in the main pillars of philosophy; philosophically qualified thinking patterns that end up in behaviour were described in Chapter 2. The status of philosophical quality may be a competence or a performance, giving rise to scores on tetralogue measurement that refer to an existing entity in our world.

Corroborating tetralogue's validity

The tetralogue as a measure of philosophical quality was designed on theoretical grounds, referring to the main pillars of philosophy and basic concepts on wisdom. Philosophical quality is assumed to be composed of five partly overlapping attributes, called indicators. These building stones were tested empirically on the possibility to be measured objectively and reliably. Computation of Cronbach's alpha-coefficient in Chapter 4 showed the homogeneity of these indicators, referring to the degree of measuring the same 'something' or quality. Subsequently, computed indices on individual (pq) and group level (PQ) data are assumed to reflect the philosophical quality of participant's thinking patterns. The validity of the tetralogue as a measure of philosophical quality was already partially corroborated in Chapter 4 through the replication of estimations of philosophically qualified tetralogues in computed PQ indices, and also through index replication of the same participants over two different tetralogues. In this chapter, empirically collected data will be examined in order to replicate theoretically based relationships between philosophical quality and other constructs as laid down in the nomological network.

2 Theoretically expected relationships between participant's characteristics and philosophical quality

A nomological network shows the assumed pattern of theoretically expected relationships between philosophical quality as central construct and several of participant's characteristics. This assumed pattern is replicated if the expected relationships are mirrored in empirical relationships, or in significant correlations between corresponding values. As explained above, this mirroring is not proof of a kind of causation but corroborates the construct validity. If positive relationships are expected and corresponding correlations are proven to be significant, they will support the claim of construct validity of the tetralogue. If a characteristic is expected to be independent of philosophical quality, and no significant correlations are found, this finding partially corroborates the validity of the tetralogue as well. If no theoretical expectations exist, potential relationships between philosophical quality and participant characteristics will be explored: their outcomes do not contribute to validation.

Figures 6.1 and 6.2 respectively present nomological networks centred on the construct of individual philosophical quality and the construct of tetralogue philosophical quality. The construct of *individual philosophical quality* must be distinguished sharply from the construct of *tetralogue's (group) philosophical quality*. Their relation is described previously in sections 4.4.4 and 4.4.5.

Philosophical quality and intelligence, educational level

For thinking patterns, a relationship between philosophical quality and intelligence is expected on theoretical grounds because both draw on analytical and reasoning qualities. It is expected with respect to individual performances and to group

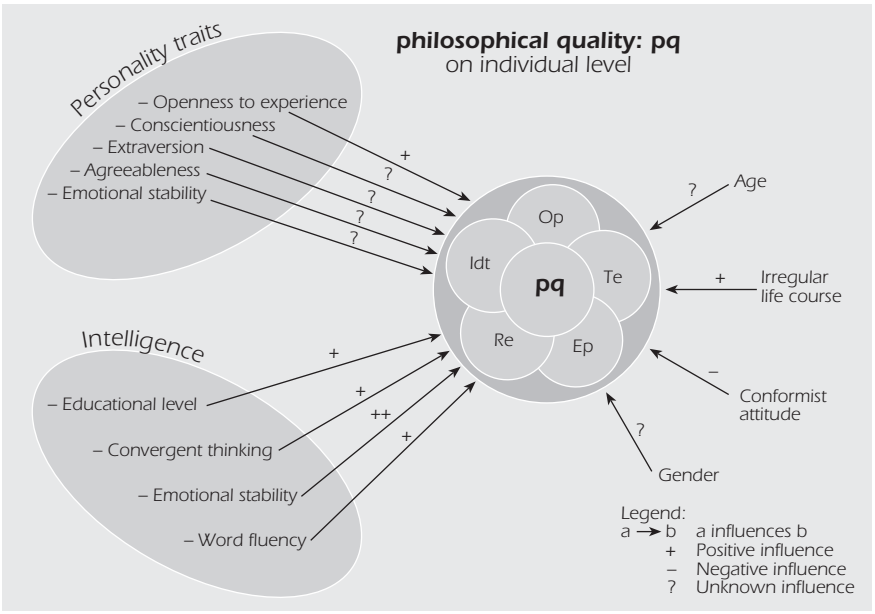

Figure 6.1. Nomological network centred on the construct individual philosophical quality.

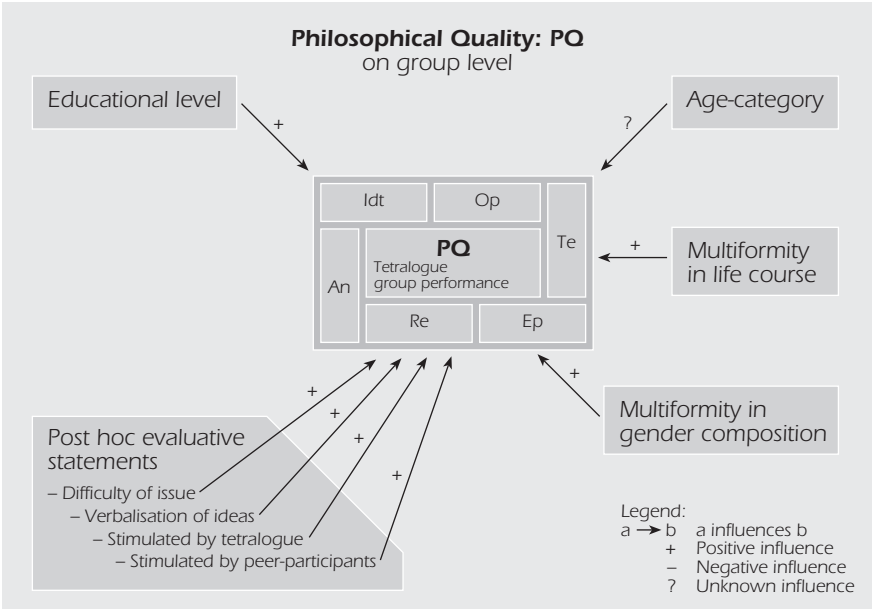

Figure 6.2. Nomological network centred on the construct tetralogue's philosophical quality.

performances. Intelligence is considered as convergent thinking: a kind of maximum performance or academic achievement for questions that comprise good answers. Convergent thinking patterns are expected to score high in intelligence tests and school achievements and will increase with age. Convergent thinking patterns can be detected through educational levels, expressed in school types, or in recommendations for school types, through tests, and through school

achievements expressed in marks. Tests and marks (GPA: Grade Point Average) present individual scores. Educational level is a characteristic of individuals, but also of groups since tetralogue groups were composed accordingly. The intelligence of youngsters is usually measured along educational lines as it is expected to 'cause' respective school achievements. Psychometric intelligence is positively correlated with educational level (Meyer et al., 2001). Performed educational levels are also the results of school selection processes. Different intellectual levels are reflected in the selection of secondary school types (e.g., VWO-HAVO: high educational level schools or VMBO: low educational level schools in The Netherlands). However, although primary school pupils are not split into their secondary school type as an indication of their level of intelligence, judgements of teachers are available. Teachers judge the pupil's intelligence as a tool for recommending appropriate secondary school types.

Convergent and divergent thinking

Intelligence is supposed to be the best measurable cognitive capacity with respect to individuals if concentrated on convergent thinking patterns. Apart from school types, two refinements of intelligence identification were used: 1) Raven-test on non-verbal behaviour, scoring fluid intelligence as pure reasoning; and 2) Grade Point Average (GPA) for languages as a proxy score for word fluency as crystallised intelligence. Philosophical quality is considered to contain both convergent and divergent thinking. Philosophically qualified thinking patterns refer to reasoning styles, including convergent thinking but emphasise divergence in thinking patterns, and more in particular, sensing ambiguities, vagueness and uncertainty, and the smooth recognition and translation of abstract patterns in concrete experience and vice versa. Philosophically qualified thinking patterns reflect questioning rather than answering; increasing complexity rather than reduction; and searching than rather finding final results. This suggests other linkages in addition to academic achievement. Sternberg (1990, 2003) points to this difference with respect to wisdom (see Chapter 2). Such tendencies may mute the effect of analysing and reasoning skills on philosophical quality. If relationships between measurements of intelligence and philosophical quality are identified, they will be subdued because philosophical quality implies mainly divergent thinking. As both attributes, philosophical quality and intelligence rely on convergent thinking only partly, the correlationship both variables is supposed to be positive, but not high. Anyhow, empirically, the data should show the distinctiveness of philosophical quality from intelligence. This requirement is met when moderate but significant correlations between empirical indices of philosophical quality and indices reflecting intelligence are found.

Collective performance

The philosophical quality of group performances is the outcome of a collectively undertaken process of inquiry. The relationship between individually performed philosophical quality and individual intelligence differs from the relationship

between collectively performed philosophical quality and the educational level of a group. Tetralogue scores reflect more than accumulated individual scores (see Chapter 4). Tetralogue scores also reflect outcomes of collective performance and dialogical events, expressing partly the intellectual character of the concerning school type. In general, it is more common in Gymnasia to question knowledge and experience than it is in schools for VMBO. A positive relation between philosophical quality and educational level is therefore expected as expressed in a significant correlation between PQ indices and school types.

Philosophical quality and personality traits

In Chapter 2, a relation was indicated between wisdom and 'openness to experiences': one of the 'Big-Five' personality traits (Staudinger et al., 1998; Brugman, 2000). In a philosophical, inquisitive and insightful sense, Saucier & Goldberg (1994) found some connection between openness to experience and autonomy, while Hendriks (1997) identifies openness to experience as 'intellectual autonomy'. Philosophical quality is assumed to be involved in the autonomous production of thoughts. The description of one of the main features of philosophical quality, sensitivity or openness to ambiguity, vagueness, and uncertainty, resembles that of the personality trait of *openness to experience*. Therefore, correlations are expected between indices for pq and scores on the factor 'openness to experience' of the NEO FFI test of personality traits. Other personality traits of the 'Big-Five' measured by this test are: *agreeableness, conscientiousness, extraversion,* and *emotional stability*. Ashton, Goldberg and Lee (2004) demonstrate that a 5-factor or a 6-factor model can represent a set of 1,710 personality-trait adjectives in seven languages. The fifth factor, referring to intellect, imagination and unconventionality (openness to experience) was less univocal in both models, but seems to be purified in the six-factor model stressing unconventionality. No hypotheses exist concerning relationships between philosophical quality and one or more of the other traits. Correlations will therefore be tested two-tailed. Finally, one may discriminate acquired attributes from personality traits as innate attributes by noticing changes in values corresponding to age. If philosophical quality belongs to a personality trait implied in the genetic heritage, only limited changes in pq indices are expected with age.

Philosophical quality and ratings on evaluative statements of group discussion

Philosophical quality was identified as a type of answer finding procedure within a group undertaking a systematic exploration of open questions through the production of autonomous lines of thought, increasing numbers of questions and uncertainties, and with a growing awareness of the degree of complexity. Philosophically qualified thinking patterns that are investigated should correspond with evaluative judgements of these aspects, providing an overview of the context-rich totality of group inquiry. Some of these characterisations, derived from the main pillars of philosophy, are not included in the calculation of PQ indices.

However, it may be possible to detect them by knowing a participant's rate of philosophical inquiry during a tetralogue. After tetralogue discussion, participants were confronted with the following issues:

1. Is the tetralogue's initial key question simpler or more complicated than you thought in advance?
2. Do you think it was difficult or easy to translate your thoughts in words?
3. Did remarks from other participants force you to think further?
4. Did the tetralogue force you to think further?

Philosophical quality is expected to relate to the judgement of the collective process of complicated philosophical inquiry. Thoughts are realised in words that offer an impetus to further thinking. The relation between philosophical quality and the facility to translate thoughts in words is not unambiguous. On the one hand, this translation must be difficult due to the impossibility to verbalise some philosophical efforts. On the other hand, expressed thoughts are the successful results of the questioned translations. Since these evaluative statements are judgements about the group process of collective answer finding procedures, a relationship is expected between answers to these questions and group performances of philosophical quality.

Philosophical quality and conformist attitude

There is one relationship assumed to be negative in the nomological network presented: Philosophical quality and a conformist attitude or dealing with guided knowledge and experience. A conformist attitude in dialogue behaviour is assumed here to be a significant contra indication for philosophically qualified thinking patterns as expressions of divergent thinking. As discussed in 'Performance and Progress in Philosophy' (Rondhuis & van der Leeuw, 2000), definite views, certainties, and an appeal to authorities all express absence of indecisive thinking patterns. Expressions of clear-cut contents suffer from a lack of openness and absence of doubt. The identification of contra indicators like thinking by yourself and the radical exclusion of outside authority's arguments were starting conditions sine qua non, in all tetralogues. Nevertheless, some negative influence on philosophical quality is assumed when participants let their thinking patterns be governed by conformism, rules and definite judgements; claiming universal validity, and forcing the non-debatable adoption of statements. Such a factor can be labelled as frequent confrontation with rules and authorities, and may be an expression of being raised following fixed rules. A negative relationship is expected between this style of child rearing and philosophical quality.

Philosophical quality and age

The age range applicable for this study is limited to students between 10 and 16 years of age. A clear relationship between philosophical quality and age during adolescence is not expected despite the complicated jumble of threads that connect

age and cognitive development with philosophical quality. Philosophical growth with age concerns increasing experience, expanding knowledge and skills with personal development and different historical influences. Youngsters acquire fresh knowledge and life experience while growing. On the one hand, philosophical quality may increase with age because age variables partly overlap with variables in intelligence, especially crystallised intelligence. The accumulation of knowledge and convergent thinking patterns develop at the expense of positive correlation with philosophical quality. In this perspective, philosophically qualified thinking patterns might grow through lifetime. On the other hand, the accumulation of experiences of failures may frustrate the production of autonomous thoughts. Continuously subordinating one's ideas to conventional thinking patterns may obstruct creativity in the production of thoughts. Many adults are inclined to defend their views too readily and to abandon tentative notions rather than to playfully entertain ambiguous or contradictory positions. Rapid technological changes and personal tragedies may also frustrate authentic performance of philosophically qualified thinking patterns. Furthermore, if philosophical quality is chiefly divergent thinking, it may 'develop' according to Riegel's dialectical train (Riegel, 1973) dependent on conjuncture of circumstances, independent of age. As a result of these partly challenging notions, relationships between philosophical quality and age will be explored as two-tailed tests without expectation.

Philosophical quality and life course

Life course is composed of experiences that can be categorised quantitatively or qualitatively. Quantitative accumulation of experiences is covered by age. Philosophical quality is not expected to increase with the number of conventional experiences. Besides quantitative growth of experience, qualitative differences due to irregular (non-normative biography) experiences are important. One of the hypotheses of this study refers to an irregular life course positively influencing philosophical quality: youngsters who think about their irregular life experiences would be better trained in, and equipped to perform philosophically qualified thinking patterns than peers with more regular life experiences. Simonton (1994) and Linley (2003) offer a second reason for this assumption. Both authors stress a positive relation between traumatic experiences and wise thinking patterns. Simonton states that creativity blooms during periods of revolution or oppression and increases after having suffered parental losses and traumatic youth. In Chapter 2, creativity was said to be a stimulating force to the performance of philosophically qualified thinking patterns. Consequently, Simonton's findings are in line with expecting a positive relationship between philosophical quality and the presence of irregular or traumatic experiences. If an irregular life course influences the performance of philosophically qualified thinking patterns positively, it is expected that corresponding pq indices will show higher values than those of regular experienced participants of the same age. This reasoning would lead also to the conclusion that philosophical quality would be acquired and might change with other relevant attributes.

Multiformity of life course within tetralogue group

Irregular life course is formulated as a characteristic of individual participants and can be considered inter-individually, varying between participants. When this characteristic is known in advance, consequences can be formulated on an individual level. Irregular life course can also be used on group level to express multiformity in life experience as a group characteristic. The intra-individual characteristic of multiformity within a group of tetralogue participants is the presence of differences in regularity of life courses. In an historical perspective, descriptions of philosophy reveal a stimulating force of being confronted with multiformity of experiences. Uncommon experience, for example in history when different cultures meet each other in wartime, or in times and places of foreign trade, will end up in rendering account for ones personal views. On a larger scale, the growth of philosophical schools can be observed in classic times (Verhoeven, 1973). It is therefore expected that PQ indices will increase when different life experiences play their part in the same tetralogue as a factor of collective performance and dialogical events. This is realised when one, two, three or all participants have irregular life courses different from each other and if deriving different points of view are mutually comprehended. It may happen that in a tetralogue assigned to regular experienced participants, one of the youngsters carries special life experiences. At a group level, multiformity of life experience was often not foreseen. It is expected that multiformity in life experience analysed on group level leads to the increase of collectively performed PQ indices.

Philosophical quality and characteristics selected for exploration

Possibly, an infinite number of characteristics may be explored for their potential relationship with philosophical quality. Some stem naturally from the procedure of tetralogue group composition, although they are not selected as carrying relevant relationships with philosophical quality. In the given nomological network, the following characteristics are not selected but may be relevant potentially: sex, geographical background, native language and ethnic background, religion, social and economical status. These characteristics can be explored systematically only if corresponding variables can be measured. This is the case with sex and may be possible with geographical background. Some characteristics are practically impossible to collect, to identify and to investigate and are outside the scope of this study: for example native language, ethnic background, religion, and social status. Other culturally conveyed baggage can also be mentioned. The main problem encountered when collecting data on these variables in a school setting is the divergence between label and content of a set of participants. For instance, in The Netherlands many old-fashioned school names may display labels as RK (Roman Catholic) or PC (Protestant Christian). Nevertheless, these schools go out for their way to enrol as many pupils as possible for economic reasons, irrespective of pupils' cultural or religious background.

Philosophical quality and gender

Differences between males and females were investigated with respect to preferences for philosophical topics. Significant differences between preferences of boys and girls were not observed. It is expected that participant's sex has no influence with respect to the performance of philosophically qualified thinking patterns. However, multiformity in a group's composition must be taken into consideration. In this study, multiformity in life course among tetralogue participants was emphasised and assumed to show a positive relationship with philosophical quality. Multiformity in tetralogue composition occurs at the group level. Gender experience may be categorised as a kind of life experience. Different sexes may refer to different life experiences and life styles. Although it was a priority for tetralogues to have a mixed composition, this requirement could not always been realised because males participated more often than females. Therefore, tetralogue groups have different sex composition: two boys and two girls, one boy and three girls, one girl and three boys, only boys, or only girls (more variation is possible with three or five participants). Scores of tetralogues with different male female ratios can be compared and their relationships with PQ indices will be explored in Section 6.5.

Nomological network based expectations

A summary of hypotheses follows.

Individually performed philosophical quality (pq) is expected to be related with:
- Intelligence: positively, but moderate
- Openness to experience as a personality trait: positively.
- Irregular life course: positively
- Conformist attitude in dealing with knowledge and experience: negatively.

Collectively realised philosophical quality in tetralogues (PQ) is expected to be related with:
- Educational level: positively
- Ratings on tetralogue's philosophical inquiry: positively
- Multiformity in life experience: positively.

The following relationships will be explored without expectations:
- Philosophical quality and age
- Philosophical quality and gender
- Philosophical quality and multiformity in tetralogue's gender composition
- Philosophical quality and conscientiousness
- Philosophical quality and extraversion
- Philosophical quality and agreeableness
- Philosophical quality and emotional stability.

3 Methods to measure and analyse

Figures 6.1 and 6.2 show networks of relationships in terms of operational measures on individual and on group level respectively.

3.1 Participants and their recruitment

Recruitment of schools and institutions

To investigate variations in philosophical quality as reflected by differences in tetralogue scores, participants were recruited from three different age groups, two levels of education, two levels of regularity in life course, sex, and geographical location. Dedicated institutions were approached to recruit youngsters with irregular life course. Once selected, participants were invited to complete tests and questionnaires on a limited number of other characteristics, like fluid intelligence and personality traits.

Table 6.a. Number of tetralogues according to age, educational level, and regularity in life course.

Age in years	**Regular life course**		**Irregular life course**	
	Educational level		**Educational level**	
	Low: VMBO	**High:VWO-HAVO**	**Low: VMBO**	**High: VWO-HAVO**
11-12-13	15	10	5	6
14-15-16	7	10	9	8
17-18-19	4	4	8	9

Schools, institutes and classes were approached and recruited from different regions of the Netherlands and in the Dutch (Flemish) speaking, Northern part of Belgium. Half of the number of schools and institutes were selected on base of their irregularity in life course of their pupils. Tetralogues were classified into 12 categories (3 * 2 * 2) (see table 6.a, also presented in Chapter 4) comprising three age levels: 1) 11 to 13 year olds, 2) 14 to 16 year olds, 3) 17 to 19-year olds; two educational levels: 1) middle or low level (VMBO), 2) high or semi-high level (VWO and HAVO); and two types of life course: 1) youngsters with regular life course, 2) those with irregular or non standard life course (physically disabled, imprisoned, or living without their biological parents). Tetralogue groups were preferably composed of participants of similar age and educational level in order to facilitate comparison of groups on these features. Irregular life course and multiformity of life experience are expected to operate respectively on individual and on group level. Tetralogues realised in institutes for irregular-experienced pupils are scored as to multiformity of experience. Mixed gender composition of tetralogue groups was aimed for, consisting preferably of two boys and two girls. More boys participated than girls. This uneven gender distribution was particularly apparent in tetralogues performed at a technical school and a prison for adolescents.

In total, 14 regular schools were recruited and 7 institutes concerning youngsters with irregular life courses. Regular schools include Primary Schools, VWO-HAVO Schools, and Schools for VMBO. In the age group of 11 to 14 year olds, some tetralogues were undertaken in Primary Schools, others in Secondary Schools for VWO-HAVO or for VMBO. Primary schools comprise five schools (15 tetralogues) with 11 to 12 year olds at educational levels estimated according to their prospects. HAVO-VWO-schools comprise five schools (22 tetralogues) with participants of the same (high and semi high) educational level, split in three age categories. Schools for VMBO comprise four schools (19 tetralogues) with participants of the same (low) educational level covering three age levels. These numbers include reserve tetralogues and exclude follow-up tetralogues with the same participants.

With respect to irregular experienced youngsters, the following institutions were approached:

- An institution for visual handicapped children, comprising four tetralogues with participants of the same age (11 to13 years old) and the same educational level (HAVO).
- A Belgian Boarding school for children and youngsters of itinerant people (shipmen and showmen), comprising nine tetralogues with participants in three age categories. Seven were conducted with low educated participants. Only two tetralogues involved higher educated participants.
- A secondary school for physically disabled youngsters, comprising 16 tetralogues with participants in three age categories and two levels of education.
- An institutions of ambulant mental care, comprising two tetralogues with the same four highly educated female participants between 14 and16 years old.
- A prison for young men, comprising seven tetralogues with participants at a low educational level in two age categories: 14 to 16 years old and 17 to 19 years old.
- Two tetralogues were conducted with adopted youngsters between 14 and 16 years old, mainly at a low educational level.
- A secondary school for youngsters with educational problems, comprising four tetralogues with low educated youngsters, 14 to 16 years old.

Occasionally, individuals scored as 'regular life course' youngsters had very special life experiences. Examples include a teenage mother and a Polish orphan living with his grandparents in The Netherlands. These tetralogues were subsequently categorised as compositionally multiform groups. Without doubt, many special experiences may be kept hidden by 'regular life course' participants, but were not detected and so tetralogues were classified as 'regular.'

Categories of individuals being difficult to approach

Contrary to the simplicity of the design, some cells of Table 6.a reflecting different combinations of participant's characteristics were hard to fill: for example irregular experienced, highly educated children and regular experienced youngsters with a low education between 17 and 19 years of age. With respect to the latter group, most are not found in schools or institutes that can be approached systematically.

Difficulties in approaching highly educated, irregularly experienced participants may arise because of the educational difficulties experienced by these participants and by educational policies in general. It must be admitted that irregular life courses involve difficulties with educational consequences. Many handicapped, blind, adopted, or imprisoned youngsters experience serious developmental problems that lead to loss of courses and resulting in lower educational achievements. This especially happens at the end of their school careers (i.e., between 17 and 19 years in age). On the other hand, at the start of a secondary school carrier, irregular experiences are often inconspicuous or have yet to be developed consciously. Developing criminality does not result in imprisonment for children aged between 11 and 13 years old. Separate treatment as consequence of irregularities in experiencing of and association with biological roots (as with adopted youngsters), usually happens during the secondary school period. Therefore, it is hard to detect irregularly experienced children and youngsters at an early stage, particularly through official lines. This is reinforced by the current educational policy of political correctness in the Netherlands that aims at integrating irregulars into regular groups without being labelled as such. These youngsters collect in specialised centres only when they drop out of conventional schools.

Demographic distribution and language

Schools and institutes were selected in different demographic areas in The Netherlands and in Belgium. Out of 95 tetralogues, 45 were conducted in the 'Randstad', a metropolis agglomerate in the central western part of The Netherlands, two in Brussels, the capital of Belgium, 34 tetralogues were conducted in midsized cities (100,000 to 200,000 inhabitants), and 14 tetralogues were conducted in provincial cities with less than 100,000 inhabitants. Demographic characteristics do not appear to be criteria for selecting familiarity with exotic cultures. Consequently, these are not expected to be relevant in this context and not explored here. Participation in tetralogues within schools and institutes was voluntary. All tetralogues were performed in Dutch, including five tetralogues conducted in Brussels with native French and Arabic participants. However, most of their utterances were unintelligible. Only two of the tetralogues in Brussels were processed. In The Netherlands, youngsters with other native languages did not participate based on their representation in school population.

As a result, in total 281 youngsters participated in 95 tetralogues: 173 males, 108 females; 147 high or semi high educated, 134 low educated; 175 regular experienced, 106 with irregular life experience. With respect to age: two participants were 10 years old, 98 participants were 11 to 13 years old, 113 participants were 14 to 16 years old, 66 participants were 17 to 19 years old, one participant was 20 and one was 23 years old.

3.2 Procedures and instruments

Data collection was done in three sessions: 1) introduction; 2) tetralogue; and 3) collecting relevant individual information. This information consists of participant's bio data and biography, and of completed intelligence test and questionnaires.

Bio data, tests and questionnaires

Completion of tests and questionnaires was scheduled after the tetralogue performance, both for one class hour (approximately 45 minutes). Duration time of test and questionnaire completion differs from participant to participant. The more serious tetralogue's participants generally took more efforts and time to complete this questionnaire than other participants. However, these effects were not taken into account in this study.

Tests and questionnaires consist of:
1. Questionnaire concerning biographical details
2. Four ratings on tetralogue's philosophical inquiry
3. Information about GPA on language performance
4. Raven test on non-verbal intelligence
5. NEO FFI questionnaire on personality traits

Table 6.b. *Descriptions of individual tests and questionnaires.*

		N Participants	Observed values (& Theoretical values)			Std. Deviation
			Minimum	Maximum	Mean	
Bio data	250					
Ratings on evaluative statements		250				
GPA languages		221	4 (0)	9.5 (10)	7.2	0.9
Raven-test C, D, E (raw)		180	1 (0)	36 (36)	23.9	7.0
NEO	Openness to experience	209	24 (12)	58 (60)	38.1	7.1
FFI	Conscientiousness	209	24 (12)	57 (60)	40.4	6.6
Rough	Extraversion	216	24 (12)	58 (60)	43.4	6.7
scores	Agreeableness	213	27 (12)	57 (60)	40.5	5.4
	Emotional stability	213	14 (12)	57 (60)	33.1	8.3

Biography and ratings on evaluative statements

Each participant is represented in a statistical file through an individual identifying number and through values corresponding to his or her characteristics. The biographical questionnaire consists of the participant's name, year of birth, school (including the educational level), class, and the date. It offers values on age in years, sex (1: male, 2: female), and school type (1: VWO-HAVO, 2: VMBO).

Ratings on evaluative statements consist of the following questions:

1. Was the tetralogue's initial key issue simpler (1) or more complicated (2) than you thought in advance?
2. Did you think it was difficult (1) or easy (2) to translate your thoughts in words?
3. Did remarks from other participants force you to think further, a little bit (1) or not very much (2)?
4. Did the tetralogue force you, yes (1) or no (2), to think further?

For each question two options were given (underlined): (1) for the first option and (2) for the second.

Intelligence and word fluency

Participant's convergent thinking (intelligence and word fluency) was measured in three ways: 1) educational level; 2) GPA verbal; and 3) results on the Raven test. The educational level is displayed nominally: a) to the higher educational level of VWO-HAVO and b) to the lower educational level of VMBO. Each participant was asked to provide his or her Grade Point Average (GPA) on language performance. This mark represents verbal school achievements on all languages the pupil had been schooled in over the last year. Although this mark reflects a judgement by the schoolteacher, the GPA verbal is used as a measure of word fluency in relation to the requirements of the school level (type and year), facilitating comparison with pupils in the same school. Participant's non-verbal intelligence was assessed by the Raven-test (Raven, 1977), consisting of five sets of 12 items each to complete. Only the sets C, D, and E were used. Sets A and B are less discriminative in the relevant age group and are therefore often neglected in testing secondary school pupils (Zeeuw, 1971).

Personality traits

To collect data on personality traits, participants were asked to complete the NEO-FFI Big Five Persoonlijkheidsvragenlijsten (Hoekstra, 1996). This consists of 60 items that can be agreed upon with on a 5-points scale (Costa & McCrae, 1989). Each item contains one of the five personality traits (12 items to indicate one trait): scoring results in five rough scores for each personality trait with values between 12 (12 * 1) and 60 (12 * 5).

Conformist attitude

Although a relation between philosophical quality and conformist attitude was indicated in the nomological network, this characteristic was ultimately neglected. Initially, the Nijmeegse Opvoedingsvragenlijst (N.O.V.) (Ammers e.a., 1998) was selected to check the degree of inclination or disposal to think and act according to rules. This questionnaire consists of 16 items with statements about the

behaviour of participant's father and mother that have to be judged on a 6-point scale. However, these items are ambiguous and formulated in terms of 'my mother/father says ...' and have to be judged in terms of agreement. Although the instruction requested the parent's judgement, but many participants presented their own judgement. It was thus not clear which part of the agreement should be attributed to the participant and what to the parent. As a consequence, results of this questionnaire had to be ignored. A systematic investigation of this factor was not possible within the time frame of the study.

As a result, out of 281 participants in 95 tetralogues, 250 youngsters returned questionnaires with bio data and evaluative statements, 221 participants offered their GPA for languages, 180 participants completed the Raven-test, and 209 participants completed the NEO FFI questionnaire on personality traits. Minimum, maximum, and mean of GPA, Raven-test results (rough scores), and NEO FFI-questionnaire results (rough scores) are presented in Table 6.b.

3.3 *Plan of analysis*

Empirical relationships between data on philosophical quality and participant's characteristics will be analysed at the individual and group level. Analyses of individual level data concern relationships between individually performed philosophical quality (pq indices) and selected characteristics of individual participants. Analyses of group level data include group characteristics expressed in modal participant characteristic or in multiformity of tetralogue's composition. These analyses cover relationships between collectively achieved philosophical quality (PQ indices) and group characteristics. Group characteristics can be imported in a set of individual data and used in analyses of inter-individual variations, but the reverse is not possible. For both kinds of information, individual and group level data will be used; first to corroborate the construct validity of the tetralogue and its associated indices pq and PQ, and secondly for explorative purposes. For corroboration, only the relationships selected in Section 6.2 will be taken into consideration; while for exploration, all measured although not selected characteristics are qualified (age, sex, conscientiousness, extraversion, agreeableness, emotional stability). Table 6.c.1 and 6.c.2 present surveys on relevant qualities or characteristics, their empirical expressions in variables and values, and expectations with respect to their relationships with philosophical quality

Relationships on individual level

On the basis of the nomological network, individually performed philosophical quality (pq indices) is expected to be related with participant's intelligence, openness to experience and irregular life course. In this study, intelligence is expressed in three variables: educational level or school type, results on the Raven test, and GPA verbal. Correlations will be computed between pq indices and values corresponding to these variables of intelligence. Correlation between pq indices and openness to experience as a personality trait and between pq indices and

Table 6.c.1. *Variables, values, and expectations with respect to relevant individual qualities.*

Characteristics of individuals	**Empirics**		**Expected relations with philosophical quality**
	Variables	**Range of values**	**pq indices (individual)**
Age	Age	10 - 23	No expectations
Intelligence or convergent thinking	School type	1: VMB 2: VWO-HAVO	Negative
	Raven-test	0 - 36	Positive, moderate
	GPA	0 - 10	Positive, moderate
Personality traits	Openness to exp.	12 - 60	Positive
	Conscientiousness	12 - 60	No expectations
	Extraversion	12 - 60	No expectations
	Agreeableness	12 - 60	No expectations
	Emotional stab.	12 - 60	No expectations
Conformist attitude	Not measured		Negative
Life course	Regular/irregular	1: regular 2: irregular	Positive
Sex	Male or female	1: male 2: female	No expectations

Table 6.c.2. *Variables, values, and expectations with respect to relevant tetralogue qualities.*

Characteristics of performed tetralogues	**Empirics**		**Expected relations with philosophical quality**
	Variables	**Range of values**	**PQ indices**
Age	Age categories	1: 10 - 13 years 2: 14 - 16 years 3: 17 - 23 years	No expectations
Educational level	School type	1: VMBO 2: VWO-HAVO	Negative
Composition of life courses	Multiformity	1: uniform 2: multiform	Positive
Gender composition	Distribution of males and females on groups	1: only boys 2: one girl, three or four boys 3: 2 boys, 2 girls 4: one boy, three girls 5: only girls	No expectations

irregular experience will be computed likewise. These relationships are expected to be significant and will be tested statistically as one-tailed tests. As mentioned above, conformist attitudes could not be measured since trustworthy data were not available. An expected negative relationship between philosophical quality and conformist attitude could therefore not be tested. Measured characteristics without expectations like age, sex, conscientiousness, extraversion, agreeableness, and emotional stability, measured by NEO FFI will be explored further in this chapter.

Relationships on tetralogue level data

PQ indices are derived from collective achievements and consist of three parts: 1) the mean pq index of the three highest scoring participants; 2) the number of qualified utterances; and 3) the number of qualified successions of two utterances (see Section 4.4.5). The first part is directly related to individual performances. The other two factors refer to tetralogue features regarding the dynamics of collective and interaction forces of participants. Some of these forces are part of the process of inquiry and of chair interventions discussed in Chapter 7. Collectively achieved philosophically qualified outcomes (PQ) are expected to be related with three characteristics: 1) educational level of school type; 2) multiformity in life experience; and 3) ratings on evaluative statements. Normally, a tetralogue group consists of participants from equal educational level. In some groups from primary schools or specialised institutes, individuals of different educational levels participated. In these cases, the dominant educational level is assigned to the tetralogue. Relations between PQ indices and educational level will be tested similar to those on individual level data. With respect to tetralogue's composition, two new variables were introduced: one concerning the multiformity of participant's life course and one concerning the male/female ratio in composition of tetralogues. The variable multiformity of life experience is scored as: 1) only regular experienced participants, and 2) the presence of at least one known participant with an irregular life course. The variable of sex composition is scored as: 1) tetralogue composed of only boys; 2) tetralogue composed of one girl and three or four boys; 3) tetralogue composed of two or three boys and two or three girls; 4) tetralogue composed of one boy and three girls; 5) tetralogue composed of only girls.

Table 6.c.3. *Variables, values, expectations with respect to ratings on evaluative statements.*

Characteristics of individuals	**Empirics**		**Expected relations with philosophical quality**
	Variables	**Range of values**	**PQ indices (group-level)**
Post hoc evaluative statements about tetralogue	Difficulty of issue	1: uncomplicated 2: complicated	Positive
	Verbalisation of ideas	1: difficult 2: easy	Positive
	Stimulated by peer-participants	1: yes 2: no	Negative
	Stimulated by tetralogue	1: yes 2: no	Negative

Ratings on evaluative statements

Expected results for ratings of evaluative statements concern individual judgements about the collectively performed tetralogue. So, individually scored values must be related collectively with performed philosophical quality. In order to relate individual scores with group scores, PQ indices (group scores) were imported as a variable in the data set for individuals. Each PQ index is the result of a joint venture and

can be assigned to each of the mostly four participants in a tetralogue. When a youngster participated in more than one tetralogue, he or she was provided with the PQ index for the first tetralogue entered. Subsequently, analysis was undertaken on the data set for individuals. Table 6.c.3 shows the design on relations between PQ indices and ratings on evaluative statements about tetralogue experience.

Corroborating and exploring

The validity of tetralogue as a measure of philosophical quality is corroborated if the expected relationships summarised in Section 6.2. and depicted in Figures 6.1 and 6.2 are replicated in empirical findings. These relationships are tested with one-tailed statistical tests. Relationships between indices of philosophical quality and characteristics measured without expectations are explored. These relationships involve age, gender, conscientiousness, extraversion, agreeableness, and emotional stability on individual level data. On group level data, tetralogue composition is involved. These relationships are tested statistically as two-tailed tests.

4 *Results related to construct validation*

Indices of philosophical quality and values of characteristics are processed statistically with the SPSS-package. Results are compared with expectations from the nomological network in order to corroborate the validation of the tetralogue as measurement for philosophical quality. Results concerning exploration of concerning relationships between non-selected characteristics are presented in Section 6.5. Results concerning relations between indices for philosophical quality and characteristics of participants, individual and collective, are presented in Tables 6.d.1, 6.d.2, and 6.d.3.

Table 6.d.1. Correlations between pq indices and selected participant characteristics.

Characteristics of participants	**Correlation with pq (individual philosophical quality)**	**Number of participants**
Educational level	0.44 **	216
GPA language performance	0.20 **	155
Raven-test set C, D, E	0.15 *	125
Openness to experience	0.27 **	149
Irregular life course	0.05 n.s.	216

Table 6.d.2. Correlations between PQ indices and selected characteristics.

Characteristics of tetralogue performance	**Correlation with PQ (tetralogue's philosophical quality)**	**Number of tetralogues**
Educational level	0.30 **	68
Multiformity in life exp.	–0.14 n.s.	68

Table 6.d.3. Correlations between PQ indices and selected characteristics.

Individually collected ratings on evaluative statements	Correlation with PQ (tetralogue's philosophical quality)	Number of participants
Difficulty of issue	0.16 *	181
Verbalisation of ideas	0.25 **	176
Stimulated by participants	–0.24 **	176
Stimulated by tetralogue	–0.16 *	181

* Significant at the 0.05 level (1-tailed); ** significant at the 0.01 level (1-tailed)

Relationships between individual characteristics and philosophical quality (pq indices)

Indices for pq were correlated with: 1) school type (1 = low level, 2 = high level); 2) results on Raven-test; 3) self reports on GPA for languages; 4) openness to experience (NEO FFI test); 5) irregularity of life course. Correlations in the expected direction are found between individually performed philosophical quality and convergent thinking as represented by participant's educational level, their GPA-verbal and their results on the Raven-test. As expected, these correlations are not too high, indicating sufficient divergence between philosophical and intelligent thinking (see Section 6.2). The correlation between pq indices and participant's general educational level is higher than that between pq indices and test results on convergent thinking. Because educational level and results on Raven-test are supposed to reflect similar capacities of convergent thinking, they are expected to correlate significantly: observed r = 0.49**. GPA's are attributed to rate the verbal capacity of pupils within a school type and cannot be used to compare pupils of different school types. This pattern of correlations clearly shows the divergence of philosophical quality with respect to intelligence and word fluency. A threshold value of convergent thinking may be involved when philosophical quality is realised. This possibility will be explored in Section 6.6.

The expected relationship between philosophical quality and openness to experience as a personality trait is confirmed by this study. Results on this personality trait measured by the NEO-FFI, also correlate with the educational level (r = 0.39**), Raven-test scores (r = 0.26**), and with GPA verbal (r = 0.18**). These empirical correlations support theoretically assumed relations between openness to experience and philosophical quality and intelligence, while at the same time showing their respective divergence. These outcomes are in concordance with the expectations. An expected relationship between pq indices and irregular life experience was not confirmed.

Relationships between characteristics and philosophical quality of tetralogue performance

PQ indices were correlated with: 1) school type (1 = low level, 2 = high level); 2) four evaluative statements (on individual level data); and 3) multiformity in life experience. As expected, significant empirical relationships are found between

philosophical quality (PQ) and educational level, also between PQ indices and four *post hoc* evaluative statements. Differences in direction of these correlations (positive or negative) are attributed to the formulation of evaluative statements (see Table 6.c.3). There is no relationship between philosophical quality and multiformity of group composition. This issue will be further discussed in Section 6.6.

Tetralogue: a valid measurement of philosophical quality?

The tetralogue was introduced to measure: 1) individually performed philosophical quality expressed in pq indices, and 2) collectively produced philosophical quality expressed in PQ indices. On both levels, several corroborative attempts were made to validate this measurement of philosophical quality. On individual level, index replication in the same participants over two different tetralogues was examined in Chapter 4. It was demonstrated that on individual level, pq index appears to be stable over different periods of time. The five indicators, constituents of the measured philosophical quality, turn out to be insensitive to random fluctuations, i.e. they are reliable. Given the high value of Cronbach's alpha, the presence of a single underlying attribute in measured indicators is plausible. As expected, low correlations, but significant, are found between philosophical quality and convergent thinking (educational level, results on Raven-test, and GPA-verbal). Also the expected correlations between philosophical quality and openness to experience are found. These relationships are an important part of the nomological network and further corroborate the construct validity of the tetralogue procedure as a measure of individual philosophical quality.

Individually performed pq indices provide a basis for collective achievements as expressed by PQ index. Their validity corroborates of the validity of the tetralogue. In Chapter 4 (Section 4.4.6) another attempt was made to validate the tetralogue as a measure of the collectively performed philosophical quality by comparing PQ indices with initial estimates of tetralogue's philosophical quality. At a group level, early estimates of tetralogue's philosophical quality appear to correlate significantly with PQ indices. As expected, correlations are found between collectively performed philosophical quality and the dominant educational level, and four *post hoc* expressed ratings on evaluative statements. Replication of these relationships further corroborates the construct validity of the tetralogue as a measure of collectively performed philosophical quality.

No relationship was found between individually performed philosophical quality and irregular life course and between collectively performed philosophical quality and multiformity of life experiences of group composition.

5 Exploring relationships between philosophical quality and participant's characteristics

No expectations exist with respect to relationships between philosophical quality and age, gender (on individual level data) or male/female ratio composition (on

tetralogue level data), conscientiousness, extraversion, agreeableness, emotional stability measured by the NEO FFI, and four *post hoc* rated and evaluated statements. These relationships are explored statistically by two-tailed tests. Correlations are presented in Table 6.e.

Table 6.e. *Observed correlations between philosophical quality (individual and group) and selected characteristics.*

Characteristics of Individuals (ind.) and tetralogue groups (gr)	**Correlation with pq**	**N Participants**	**Correlation with PQ**	**N Tetralogues**
Age (ind. & gr.)	0.07 n.s.	216	0.06 n.s.	68
Gender (ind.)	0.13 n.s.	216	-	
Gender composition (gr.)	-		0.13 n.s.	68
Conscientiousness (ind.)	–0.05 n.s.	146	-	
Extraversion (ind.)	0.02 n.s.	153	-	
Agreeableness (ind.)	0.18*	152	-	
Emotional stability (ind.)	–0.05 n.s.	150	-	
Difficulty of issue (ind.)	0.12 n.s.	182	See Table 6.d.3	
Verbalisation of ideas (ind.)	0.23**	177	See Table 6.d.3	
Stimulated by participants (ind.)	–0.01 n.s.	177	See Table 6.d.3	
Stimulated by tetralogue (ind.)	0.00 n.s.	182	See Table 6.d.3	

Legend: * Significant at the 0.05 level (2-tailed); ** significant at the 0.01 level (2-tailed)
pq: individually measured philosophical quality
PQ: philosophical quality measured for tetralogues
ind: individually
gr: group

The gender composition of a tetralogue was categorised in five values: 1) tetralogues composed of only boys (N = 22); 2) tetralogues composed of one girl and three or four boys (N = 10); 3) tetralogues composed of two or three boys and two or three girls (N = 24); 4) tetralogues composed of one boy and three girls (N = 8); 5) tetralogues composed of only girls (N = 4). There was no significant relation between individually performed philosophical quality and participant's sex, or between collectively performed philosophical quality and male/female ratio composition. A slightly significant correlation is found between individually performed philosophical quality and agreeableness as personality trait, and also between individually performed philosophical quality and participant's estimation of their capacity to verbalise ideas. Although this issue concerns an activity in a collective process, it is the only out of four evaluative statements that stresses an individual capacity, so this relation is not surprising.

6 *Discussion about potential effects of age, gender, irregular life course, and a threshold value of convergent thinking*

Investigation and exploration of expected and unexpected relationships offer some answers, but raise other issues as well. Such questions and issues involve effects

of age, gender, and irregular life course on philosophical quality and the existence of a threshold value of convergent thinking to perform philosophical quality. Although a correlation between age and philosophical quality was not found, this characteristic discussed heavily (see Chapter 2) has potential as an influencing factor on wisdom. Differences are experienced in the performance of philosophically qualified thinking patterns between youngsters at primary schools and those in secondary schools. In particular, the effect of the age grade of school transition on philosophical quality requires further discussion and research. Although gender composition of tetralogue groups did not influence philosophical quality, differences between the sexes were observed, especially with respect to the performances of two indicators: *ep* (epistemological position) and *an* (anecdotal quality) concerning utterances with personally loaded expressions. Finally, the effects of irregular life course and multiformity of life experience on performed philosophical qualities in tetralogue composition are discussed as the results were not in concordance with expectations.

Age

Age effects on individually performed philosophical quality and tetralogue's philosophical quality were not found. Relationships between philosophical quality and age or the accumulation of regular experiences were not expected although extensively debated. Opposite views concerning this topic were summarised in Section 6.2. The fact that there was no correlation between age and philosophical quality at individual and tetralogue levels may indicate the insensitivity of philosophical quality to the maturation process. This result clearly differs from most findings concerning acquired attributes. It is possible that philosophical quality cannot be acquired and may be part of innate personality traits.

Shifting pq's after school transition

Differences were noticed in the performance of philosophical qualified thinking patterns between pupils from primary schools and pupils from secondary schools. Therefore, a division was made between two groups: 1) participants from primary schools (N = 59), and 2) participants from secondary schools (number of participants with investigated pq indices: N = 182; total number: N = 249). Testing for relationships between the two groups produces different results: correlations between pq's and measures for convergent thinking (educational level, GPA verbal, Raven results) among pupils from secondary schools are respectively: $r = 0.51^{**}$, $r = 0.25^{**}$, $r = 0.21^{*}$, while no significant correlations were found among primary school pupils at all. The same tendency holds for correlations between individually performed philosophical quality and openness to experience as a personality trait and also to correlations between pq's and the results on four evaluative statements. While no significant correlations were found among primary school pupils, corresponding correlations among participants of secondary schools are: with respect to openness to experience $r = 0.33^{**}$, and to the four evaluative statements respectively: $r = 0.21^{**}$, $r = 0.27^{**}$, $r = -0.24^{**}$, $r = -0.23^{**}$. However, tetralogues

performed on primary schools are shown to be reliable measures (see Chapter 4). No reason could be found to question the correct implementation, reliability, and validity of measurements of participant's characteristics and of philosophical quality at all ages. This suggests that unidentified factors may play a role.

Individual philosophical quality, measured through tetralogues in primary schools, seems to be distributed equally among children of different intellectual levels. Contrary to this result, youngsters at grammar school show significantly higher pq indices than their peers on technical and vocational training schools. This effect converges with the age grade of transition from primary to secondary school. However, age cannot be identified as an influencing factor on philosophical quality after entering secondary schools. Is the measured difference in philosophical quality between schools of high and low educational level a consequence of increasing pq indices at VWO-HAVO schools or of declining pq indices after entering VMBO schools? The intellectual climate at a grammar school is certainly more likely to encourage uncommon questions from students than in modern technical and vocational training schools. After all, the possibility to raise uncommon questions is a primary condition to score philosophical quality in tetralogues. The intellectual climate of secondary schools is the result of many factors, including characteristics of scholars, their background and educators, school programs and political decisions. Further investigation of these factors are beyond the scope of this research. Differences in indicator frequencies between primary and secondary school pupils may provide some insight. There is a significant difference between primary and secondary school pupils in their use of expressions with *ep* (epistemic position). Participants in primary schools are significantly less inclined to express themselves from an epistemic position than pupils of secondary schools: $r = 0.24^{**}$ ($N = 216$). A similar correlation is computed for group level data collectively scored as indicator frequencies summarised for each tetralogue. The relationship with respect to *ep* scores is even more striking, with $r = 0.43^{**}$ ($N = 70$).

The question of shifting philosophical qualities after leaving a primary school can and will be investigated by following children during their transition from primary to secondary school. A group of four primary school boys was followed in nine tetralogues during a period of 28 months as they moved from primary school to VMBO schools. Description and results of this investigation will be presented in Chapter 8.

Gender and gender composition of a tetralogue group

Although gender was not expected to effect philosophical quality, differences between participants in expressing themselves were remarkable. Girls seem to provide ego-involved statements, using I, noticeably more than boys who tend to express themselves in general terms. One of the indicators constructing pq indices is *ep*: epistemic position, which is an indicator of cognitive expressions referring to *I*. It is possible that this verbal behaviour of expression using ego-involved terms influences the individually performed philosophical quality. A comparison of mean *ep* frequencies is made between boys and girls, and demonstrates a

significant difference. Correlation between *ep* frequencies and sex: $r = 0.15^*$, $p < 0.05$. The same holds with respect to indicator frequencies of *an* (anecdotal quality), pointing to real-life experiences and descriptions of concrete situations. Comparing the means of *an* frequencies between boys and girls also demonstrates a significant difference favouring girls' use of anecdotes in contributions. The correlation between *an* frequencies and sex shows that $r = 0.14^*$, $p < 0.05$. Notwithstanding these results, significant correlations between individually performed philosophical quality and gender, or between a tetralogue's philosophical quality and sex, or male/female composition are not observed.

Irregular life course and multiformity in life experience

Although expected, relationships between individually performed philosophical quality (pq indices) and individual irregular life course; and between collectively performed philosophical quality (PQ indices) and multiformity in life experience in tetralogue groups were not found. However, the relationship between multiformity on life experience and PQ indices appears to be influenced by the educational level as can be shown by the result of variance analysis of PQ indices for these two factors (multiformity * educational level): $F(1, 64) = 5.88$, $p = 0.02$. Subsequently, in the data set of group characteristics, correlations were computed between PQ indices and multiformity in tetralogue composition, separately for high educated and low educated groups. Correlation coefficients are respectively: $r = -0.43^*$ ($N = 34$), $r = 0.13$ n.s. ($N = 34$). So, only within highly educated groups does the collectively performed philosophical quality appear to be significantly higher than within the lower educated groups, but only if no participant is identified with an irregular life course. For the data set on individual characteristics, the same division into two levels of education was made. Here, no correlations were found among highly educated participants. However, among low educated participants, a significant correlation between individually (pq) and collectively (PQ) performed philosophical quality with life experience is observed, respectively: $r = 0.24^*$ ($N=105$) and $r = 0.39^{**}$ ($N=104$). Contrary to the results for group level data, individual philosophical qualities of participants with an irregular life course turn out to be significantly higher than those of participants with a regular life course, but only within low educated groups. This result, concerning low educated participants, measured on individual level data is in concordance with the expectations. Other results from this section are difficult to explain and are outside the scope of this study.

Regular and irregular life course

The characteristic of regular or irregular life course was incorporated into this research on the assumption that irregularities in life course lead to individual pondering and to qualified exchange of thinking patterns at the tetralogue level. However, quantification of regular and irregular life courses is questionable. In this study, *irregular life course* encompasses blindness, living without (biological) parents, or imprisonment. A *regular life course* was not defined and not measured.

It is even possible that many regular participants may have had non-regular life experiences in their lifetime: a value *regular life course* was only assumed. Many experiences considered to be non-regular remain hidden within the characteristics of regular participants, for example when youngsters assigned to regular life course demonstrate special experiences like teenager motherhood or being a refugee. These occurrences are telling and indicate a potential ubiquity of irregular experiences amongst participants in tetralogue groups. Consequently, the value of a *regular life course* may not be appropriate for identifying significant relationships between philosophical quality and irregular life course, or between a tetralogue's philosophical quality and the multiformity of it's group composition.

Threshold value of intelligence

It is likely that there is a threshold value of intelligence for the philosophical quality of a performance. If a threshold value exists, philosophical quality indices will increase significantly with rising intelligence values only beneath or above this boundary. It is expected that this happens only beneath such threshold. At the other side of the threshold, no significant correlation between philosophical and intelligence qualities will be found. To explore this possibility, participants were divided in two levels of intelligence. In each group, a correlation will be computed between indices of philosophical quality and those of intelligence, independent of the dividing criterion. In this study, three independently measured variables are available for detecting levels of intelligence: educational level, results on Raven test, and GPA verbal. From these, GPA verbal is measured according to standards of the respective schools and is, therefore, theoretically unrelated to the educational level. Correlation results according to the group divisions are presented in Tables 6.f.1, 6.f.2, and 6.f.3.

When classifying participants according to their educational level, no differences are observed between correlations of pq indices with indices for convergent thinking at low and high educational levels. Classifying participants according to their GPA-verbal below the value of 8 or >= 8, clearly shows a threshold value of verbal intelligence with respect to the relationship between philosophical quality and educational level. Below this boundary, significant correlations between indices for philosophical quality and educational level are seen. Above this threshold significant correlations disappear, as is expected. So, a threshold value of verbal intelligence is plausible. When classifying participants according to their Raven results, a significant division criterion is found at 23 (number of correct answers), but only with respect to the relation between indices for philosophical quality (pq and PQ) and GPA's. If participant score beneath this boundary of 23 correct answers, their philosophical quality (pq and PQ) is not related with their GPA verbal, while above a Raven score 23, the philosophical quality significantly rises with GPA verbal. This result contrasts with the result concerning a division according to a GPA threshold and is difficult to interprete because GPA's represent dividing criteria within the same educational level only.

Table 6.f.1. Correlations separately measured for low and high achievements on educational level.

Low educ.	pq indices	N Part.	PQ indices	N Part.	High educ.	pq indices	N Part.	PQ indices	N Part.
GPA-verbal	0.11 n.s.	69	0.13 n.s.	68	GPA-verbal	0.14 n.s.	86	0.00 n.s.	86
Raven	0.05 n.s.	66	0.13 n.s.	65	Raven	-0.18 n.s.	59	-0.17 n.s.	59

Table 6.f.2. Correlations separately measured for low and high achievements on GPA-verbal.

GPA < 8	pq indices	N Part.	PQ indices	N Part.	GPA >= 8	pq indices	N Part.	PQ indices	N Part.
Educ. level	0.44**	92	0.47**	110	Educ. Level	0.29 n.s.	44	0.27 n.s.	44
Raven	–0.03 n.s.	111	0.23*	91	Raven	0.40 n.s.	24	0.21 n.s.	24

Table 6.f.3. Correlations separately measured for low and high achievements on Raven test.

Raven < 23	pq indices	N Part.	PQ indices	N Part.	Raven >= 23	pq indices	N Part.	PQ indices	N Part.
Educ. level	0.52**	48	0.52**	48	Educ. Level	0.30**	77	0.48**	76
GPA-verbal	–0.03 n.s.	42	0.14 n.s.	42	GPA-verbal	0.40**	74	0.28**	73

Legend: Educ. Level: Educational level
Low educ.: Low educational level: VMBO
High educ.: High educational level: HAVO-VWO
GPA verbal: Grade Point Average for languages
Raven: Results Raven test on non verbal intelligence test

7 Summary and Conclusion

In this Chapter, characteristics of participants were explored to discover relationships at philosophical quality, performed individually or produced in a tetralogue group, and to further corroborate the construct validity of tetralogue as measurement of philosophical quality. Figures 6.3 and 6.4 show the nomological network of figures 6.1 and 6.2 provided with results of empirically replicated relationships at individual and group levels in this study.

The construct validity of a tetralogue as a measurement device for philosophical quality is plausible because most of the expected relationships from the nomological network(s) are replicated. Indices of pq and PQ are valid measures of the philosophical quality of thinking patterns.

Philosophical quality on the individual level is significantly related to the personality trait openness to experience, to educational level, and to intelligence, although to a limited extent, and is unrelated to age and gender. A complicated network of many interfering factors has emerged with respect to influences of participant characteristics on philosophical quality. Variance of age is affected by transition from primary to secondary school; the relation between philosophical quality and a supposed irregular life course is affected by the educational level and also

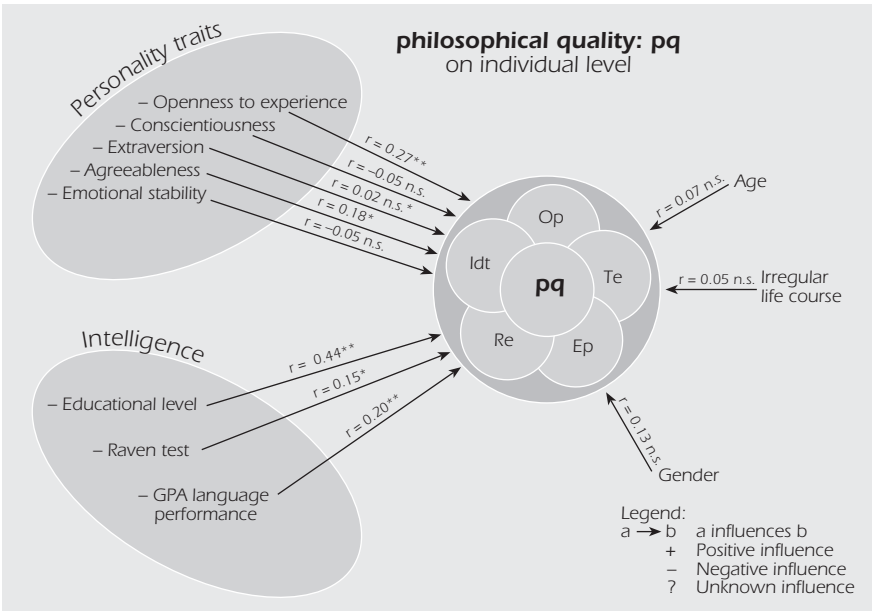

Figure 6.3. Empirically replicated nomological network of the individual philosophical quality.

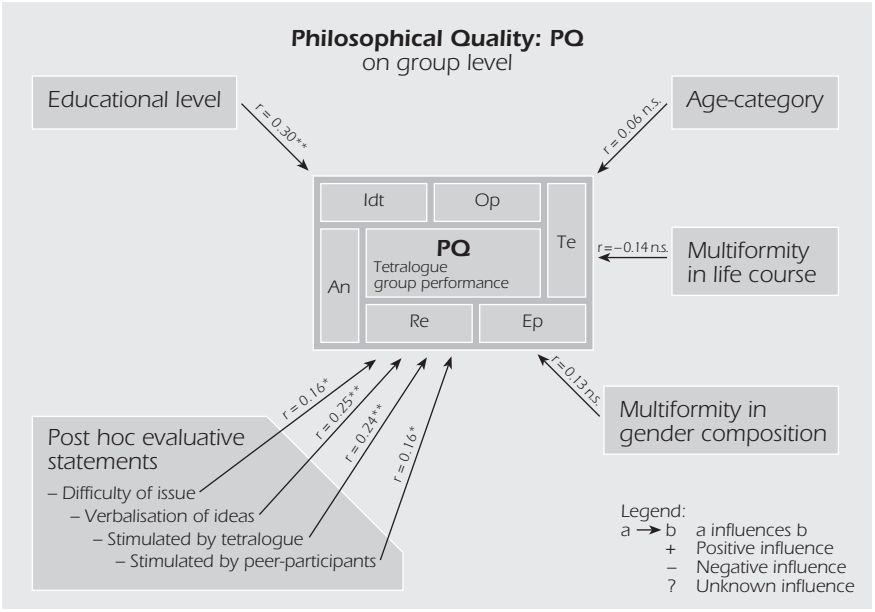

Figure 6.4. Empirically replicated nomological network of tetralogue's philosophical quality.

depends on a clear definition of a non-irregular life course; relationships between philosophical quality and educational level are affected by a threshold value of verbal intelligence (GPA verbal). A better understanding of these interferences would require further and more detailed research with respect to the characteristics mentioned and the transition period between primary to secondary school.

7 Chair Interventions and General Didactics of Tetralogues

Similarities and differences between the philosophical qualities of tetralogues are observed and investigated with respect to themes and characteristics of participants. Here, differences will be investigated from a formal point of view. The design and realisation of a tetralogue may vary according to rules, styles of being conducted and number of chair interventions. Types of philosophical discussion are described in Section 7.1. A tetralogue's philosophical character is controlled by discussion rules and chair performance. Didactic rules, conducting styles and chair interventions will be discussed in Section 7.2. Tetralogues will be explored as result of differences in conduction styles (Section 7.3) and number of interventions (Section7.4) as potentially influencing philosophical quality.

1 Philosophical discussions

Discussions

Philosophical discussions or dialogues (the words will be used interchangeably) are distinguished from other types of discussion with respect to their content, their use of thinking patterns only, and their openness to receding definite answers. Discussions in general are often found in educational settings. Educational discussions especially between students and teachers differ from philosophical dialogues in several aspects: starting issue is guided by correctness, the goal is mostly instructional, persuasive, or a justification. Sometimes the social nature of interaction is the central focus. Un-intentionally, the environment may influence the quality of discussion (Vygotsky, 1964). Contrary to discussion between student and teacher, interaction between peers facilitates making discoveries since pupils have to initiate the course of the discussion by themselves in the absence of a steering chair and requirements of correctness (Elbers, 1992).

In science, dialogues appear only in order to find a solution; moreover scientific dialogues need more than thinking patterns only: specialised knowledge and observation of reality are required. Interviews, apologias and debates are more illustrations of discussion types. Many educational institutes offer debating classes so that students may practise rhetoric skills. Although, rhetoric is part of philosophy, it does not cover all philosophical qualities as was explained in Chapter 2.
The main difference between tetralogues and debating sessions is that the latter aims at persuasion. In case of debating, discussions develop by an increasing number of established opinions and a decreasing number of questions.

Another type of discussion may be distinguished, as demonstrated in 'make-believe' play and in drama. Vygotsky (1964) emphasised the relevance of 'make-believe' play in development. In creating imaginary situations sayings must be in accord with internal ideas and satisfy the rules of the scene. Drama discussions generate

a paradox: acting in accord with rules and internal ideas realised in advance versus making believe that sayings are authentic to the audience at the same time. Drama players operate on a theatrical or meta-level, realising the transfer of their search for truth on the audience. Drama discussions may express philosophical qualities. However, this type is excluded from this study.

Philosophical discussion

Oral and written philosophical dialogues are relevant expressions of philosophically qualified thinking patterns. In classical times, truth was pursued disinterestedly through analytical discussion and by developing concepts and thoughts. Through history, periods of different philosophical approaches alternate, but yet remained linked to different philosophical attitudes. For example, in Europe, periods of Cartesian philosophy and German idealism alternate with periods of sliding concepts of reality and nebulous discipline boundaries. These latter intervals are characterised by a significantly greater number of written dialogues expressing sensitivity to ambiguities, vagueness and uncertainty within philosophy but also on a meta-level. These dialogues can be found in Greek thought during the fourth century B.C., the Italian Renaissance, and the eighteenth century Enlightenment (cf. Berkeley, Hume, Diderot). Outside Europe, the Indian *Upanishads*, the *Analects* of Confucius and the *Mencius* or *Mengzi* are examples of Eastern literary philosophical dialogues (van der Leeuw, 2005). Modern philosophical dialogue forms are: *Socratic discussions* (i.e. specialised sessions), *philosophy for children*, and *philosophical counselling*. All appear to be Socratic discussions in nature.

Tetralogues are Socratic dialogues between four participants. Although the term 'dialogue' refers to the number of two, Socratic dialogues in the work of Plato took often place with more than two participants. Accepting the Socratic method of counselling, tetralogues begin with open (philosophical) questions, focusing on themes and exploring truth as an ultimate and ever receding goal. Performing force is the community of inquiry. Ideas, arguments, analyses of concepts, and common experience are the only means to attain this goal. In philosophical discussions no special observation of reality is necessary to find solutions. All kinds of specialised knowledge, privileged access to philosophical truth or insight, intuition, mystical union are excluded. Contrarily to science the performance of thinking patterns is more important than the result. A significant consequence of this method is the suspension of judgement in order to create judgement. The Socratic method is a common and never ending attempt to reach the right path for exploring reality, starting from experience and looking for general underlying principles. At the end of the discussion, the initial question often returns in a changed form, charged with a number of different views and secondary questions, coloured by ambiguity, vagueness and uncertainty. Reality and everyday experience appear to root in a conceptual framework. Philosophical thinking can only clarify this conceptual framework, trying to understand things in cohesion with the world, with the '*eye* on the whole' (Sellars, 1963; Philipse, 2004).

2 Didactics of tetralogues

Discussion rules and chair

Socratic dialogues are characterised by openness with regard to content, and by sensible didactic rules in order to attain their goal. Didactics cover formal aspects of discussion with respect to its conduct. A chairperson formulates discussion rules settled and monitored tetralogues to prevent potential risks. Tetralogues produce autonomous and consistent thoughts through analyses of an initial question and related concepts. It is best served by an open approach of themes and by the equivalence of all participants. There is a delicate balance between the search for paradoxes and ambiguities, and preventing debate. There are many other potential risks to avoid, such as: obscurity of sayings and arguments; hidden presuppositions; rigid opinions (values); definitive judgements; verbalisms or clichés; unlimited babbling or endless repetition of arguments; intimidation of participants; and moral or political correctness.

Discussion rules concern linguistic and non-linguistic behaviour:
1. Opinions are only allowed when backed up by arguments.
2. Participants may attack arguments, but not opinions of others.
3. Sayings and arguments must be understandable, accessible and controllable to all; they must be free from hidden presuppositions or introspection.
4. Dogmas, irrational certainties, arguments based on external authority or definitive judgements are not allowed since these are not considered intelligible arguments.
5. Trains of thought should be consistent and constructed systematically.
6. Participants are permitted to interrupt unannounced, as this may encourage spontaneous expression of ideas.

With respect to the chair, several persons were invited to perform and rules were laid down for interventions. Except for the author, chair-persons from outside were selected because of their philosophical and didactical qualities: philosophy students attending didactic classes, and psychology students interested in philosophy. All received instructions prior to the sessions (see Section 4.2.2). In general, the chair's attitude was open, wondering, stimulating, and they were encouraged to prevent black-white opposition of views that may run into debating sessions. Chairs were not allowed to correct participants and emphasised that 'nobody knows the definite answer'.

Interventions

Chair interventions were qualitatively restricted to three types:
1. questioning for clarification
2. reiterating of previous remarks and questions
3. quoting a fresh illustration.

Reiteration was suggested when part of philosophical question was not discussed, when the discussion drifted too far from the selected topic, or to confront statements with earlier, and possibly contrasting ones. This last type of reiteration often proved rather constructive and provided a steering force. Offering a fresh anecdote as illustration of the initial question occurred only in very rare cases.

Chairs

Most tetralogues took place in school settings and were educationally coloured. Chair-persons were assigned to encourage the tetralogue with regards content. Sometimes they intervene to maintain order. Experience with classes and youngsters and being used to youngster's linguistic usage was helpful. Differences between conducting styles become known with respect to the selection of questions and of reiterated notions; secondly with respect to letting participants finish their utterances; and thirdly to the number of interventions. Selection of reiterated notions was based on judgement of relevance and on recognition of philosophical topics. Philosophy students were better equipped to recognise philosophical topics than their psychology colleagues, while the latter were more skilled in letting students finish. So, it was expected that philosophy students would intervene more than psychology student chairs. As explained in Chapter 3, one philosophical discussion usually generates several themes. It was up to the chair to decide when and what theme was to be emphasised by reiteration, when and which illustrations evoked the weird and ambiguous, and how clarifying questions were formulated. Differences in conducting styles will unarguably appear and they may be relevant in a tetralogue's development. Several conducting styles attempted to remove these differences by restricting interventions qualitatively to only three types. By inviting other chairs to participate, individualised conducting styles were prevented. Chair types and number of interventions were collected in the tetralogue database. Qualitative differences tetralogue performances were not scored.

3 *Differences between tetralogues related to chairs*

Differences between chairs with a philosophical and those with psychological background are noticeable. The involvement of philosophy students in philosophical question is greater than that of psychology students. Also, they use more interventions within the given constraints. Chairs with a psychological background intervene less and focus more on individuals. As expected, and based on a total number of 70 tetralogues, those conducted by philosophy students show a greater number of interventions than those conducted by students psychology, their mean number of interventions is respectively: 43 and 14; although the number of psychologically educated chairs was very small. Out of all 95 tetralogues, 75 were conducted by the author; 17 by students in philosophy, two by students in psychology; and one by a regular schoolteacher. The teacher performed the highest number of interventions at 90. Exploring the relationships between chair types and tetralogue characteristics such as philosophical topic, number of utterances (qualified and non-qualified), and most of the indicator frequencies did

not produce significant correlations. Slightly significant relationships were found between the philosophical expertise of the chair and two indicator frequencies as collectively performed in a tetralogue: *ep* (epistemological position) and *an* (anecdotal quality), respectively 0.27* and 0.25*, significant at the 0.05 level. This result means that ranging from the author as chair, to philosophy students as chairs, to psychology students as chairs, frequencies in the use of epistemological position and anecdotal quality increase. Differences in styles of chairing did not result in significantly different indices for PQ: analysis of variance comparing the three types of chair, shows $F(2, 65) = 2.7$ ($p = 0.08$). So, it is concluded that philosophical quality of a tetralogue does not differ significantly for the three types of chair in this study.

4 Differences between tetralogues related to numbers of interventions

The 70 tetralogues investigated differ in their number of interventions from 1 to 90, with a mean number of 37, and a standard deviation of 17.4. There were no significant correlations between the number of interventions and a tetralogue's characteristics, including age of participants, level of education, regular or irregular life course, and number of (qualified and non-qualified) utterances. The mean number of interventions of tetralogues with epistemological or metaphysical topics was 43; with anthropological topics, 37; with ethical questions, 25; and with questions concerning demarcation problems, 32. No significant differences in number of interventions between the four topics were found. Analysis of variance shows $F(3, 66) = 2.4$ ($p = 0.08$). Exploring relationships between number of interventions and collectively in a tetralogue, performed indicator frequencies shows only a moderately significant correlation with respect to *op*, (openness): $r = -0.30$*, significant at the 0.05 level. This suggests that the production of utterances scored with openness decreases as chairs intervene more. No significant correlation can be detected between the number of intervention and indices for PQ, so the number of interventions does not influence philosophical quality.

5 Conclusions

Tetralogues were conducted by chairs with philosophy-, psychology- and education background, and show different degrees of intervention. Although different conducting styles can be noticed, this circumstance did not lead to significant differences in the philosophical quality of tetralogues. Indicator frequencies of epistemological position and of anecdotal quality appear to increase moderately when a tetralogue's chair is less philosophically experienced, while frequencies of openness decrease when the discussion as the number of interventions increase. These outcomes imply some qualitative differences in conducting styles dependent of the chair's philosophical experience. However, there are no clear indications of significantly different PQ indices. A more systematic investigation should examine this issue more closely.

8 Follow-up

A follow-up series of tetralogues was designed because some children show remarkable changes in their philosophical interest and performance after entering a secondary school of a lower educational level even though a relationship between philosophical quality and age could not be detected. The present research aims to describe changes over a two-year period in which a school transition takes place, and secondly to replicate relative frequencies of five separate philosophical indicators on individual, longitudinal level (relevant previous results were reported in Chapter 4). Considerations inducing this study are described in Section 8.1, the selection of pupils for this follow-up series in Section 8.2. Nine longitudinal tetralogues were successively recorded over a period of more than two years during transition from primary to secondary school. Realisation and actual development of these tetralogues, including the behaviour of participants are characterised in Section 8.3. Data of these follow-up tetralogues are explored and the results presented in Section 8.4. In the last section of this chapter, conclusions and observed changes in the performance of philosophically qualified thinking patterns are discussed in the light of expectations, theory, and environmental influences.

1 Observing children in their transition from primary to secondary school

It is often emphasised that school children develop relatively uneventfully during their primary school period. In contrast, adolescents seem to experience a period of rebellion against social boundaries after having entered secondary schools. Transition from primary to secondary school involve shifts in school environment, daily schedules, peer companions, and academic challenges. Moving from the intimacy of a primary school classroom to a much larger, impersonal secondary school is stressful, as a more complex social world is entered. Simultaneously, grade point averages often drop and feelings of anonymity increase after school transition (Berk, 1997).

Observed differences between primary school and secondary school pupils

Thinking patterns of primary school children in this project demonstrate a spectrum of autonomously produced thoughts, irrespective of their intellectual level (Chapter 6). This observation is all the more remarkable as observations in secondary schools show significant differences in the performance of philosophically qualified thinking patterns between pupils of high and low intellectual levels. Children appear to often come up with concrete events in philosophical discussions. It is also emphasised that primary school children show more interest in metaphysical and epistemological discussions in comparison with secondary school pupils (Chapter 5).

References to real life evidences and anthropological topics

Theoretical approaches in children's thinking on reality increase after leaving primary school; they show up as objects of children's thought rather than mirror images of reality (Berk, 1997), and become coordinated with real life evidence. Direct references to real life evidence decreases. In middle childhood, thinking about the world and the self, shifts from concrete events towards underlying circumstances in the world and personality. Older children increasingly emphasise competences and psychological attributes because they are better equipped in coordinating multiple aspects of a situation and in reasoning about the physical world. Before entering a secondary school, children between the ages of 11 and 12 are especially fascinated by the physical world. A check on a primary school in two classes including 50 pupils together shows 26 children reporting they frequently look at programs on the Discovery Channel while watching TV. Primary school children involved in this study enjoyed discussing, for example the end of the universe, while adolescents preferred topics on people. This tendency was emphasised in Chapter 5, demonstrating a shift in topic selection from metaphysical and epistemological themes among 11 to13 year olds to anthropological questions popular among adolescents of 14 to 16 year olds. There is a tendency to extend thinking about the self toward a quest for identity. Adolescents tend to define who they are, what they believe, value, and choose to pursue in life. Teenagers in complex societies often experience an identity crisis as they experiment with alternatives before settling on a set of values and goals (Erikson, 1968). These observations and considerations suggest a decrease in the use of anecdotal quality in utterances.

Environmental changes

School transition is an environmental change. The performance of thinking patterns results partly from acting on the environment (Vygotsky, 1964). So, changes in school environment often turn out to be a determinant of changes in philosophically qualified thinking patterns. All children around the age of 12 years experience a change in school environment. However, this school transition differs between children going from primary to secondary, VWO-HAVO school and children going from primary to secondary VMBO school. Here, the focus is on pupils changing from primary school to secondary school of low educational level (VMBO).

The performance of thinking patterns is guided by the possibility offered by a situation to act and will depend on so-called 'affordances': i.e., opportunities of the environment to exploit potential qualities of the individual by guiding, channelling or restricting them. A general biological phenomenon is stressed: a seed of a violet (Viola vulgaris)planted in an Amsterdam back garden will develop into a flower that is shaped differently to the same type of plant grown in an alpine meadow: the latter tends to be small and thickset. Different environments offer a range of different qualities for nourishing and supporting the seeds. According to Jackson (1995), such a mechanism occurs in the development of adolescents as well. The socio-geographical or cultural-organisational environment offers opportunities

and encouragement to explore and define ones identity. Affordances in school environment emerge clearly for adolescents. They are realised in consistent sets of reactions, verbal or non-verbal, through the behaviour of peers and educators. The VMBO environment (low educational level) creates different affordances in comparison with grammar schools (high educational level). In both environments, young people try to achieve knowledge in collaboration with their environment. Pupils start a task with different understandings in a group of participants and arrive at a shared understanding (Vygotsky, 1964). The degree of shared understanding in a group of VMBO pupils differs from that in a group of VWO pupils. Observations of discussions in VMBO schools make the assumption of an environment in which social processes are going on, with a negative effect on philosophical quality. Philosophical approaches to reality, like discussing the realness of our world – maybe we live in a dream – are not considered to be 'cool' in VMBO schools, whereas pupils on grammar schools enjoy such discussions.

2 Selection of follow-up participants

In order to sense and explore changes in philosophically qualified thinking patterns performed by pupils in their transition period, a limited follow-up study was designed with four children. This project aimed to explore the concerns of children entering secondary schools of low educational level. During my years of practicing philosophising with pupils in primary schools (final class), I noticed that some children who were to enter schools at lower educational levels participated energetically, while others who were to enter a grammar school at the end of the year, did not want to participate in weekly philosophical meetings. The class teacher attending the philosophical sessions also observed this.

In May 2000, four 12 year-old boys were selected from Class 8 of a primary school in Haarlem (The Netherlands). These four boys, recruited from pupils with low educational prospects, were selected because of their enthusiasm. It was two months before summer holidays and before dispersing into different secondary schools. Steve, Erny, Tom and Marc (names are changed) were willing to participate repeatedly in philosophical meetings (tetralogues) over a longer stretch of time. Their autonomous production of authentic thoughts was the more remarkable as their academic achievements were of low quality according to their teacher. So, the enterprise to conduct a series of a follow-up tetralogues with this particular four was captivating and promising in the eyes of the teacher and author. The original design was to conduct some nine tetralogues during a one and a half year period. For the four students, changes in their philosophical behaviour from the age of 12 to 14 years could be sensed and explored over the period of transition from primary school to low level secondary school.

3 Nine successive tetralogues

The boys participated in nine consecutive tetralogues over a period of 28 months. Table 8.a presents an overview. Tetralogue 1 and 2 were videotaped in the last

part of the last primary school year before holidays, in primary school time. Tetralogues 3, 4, 5, 6, and 7 were videotaped during the first year of the secondary school. Tetralogue 8 in the second year of the secondary school, and Tetralogue 9 at the very beginning of the third year of attending the secondary school. Most tetralogues took place under the roof of the rooting primary school, Tetralogue 8 in the house of the author and Tetralogue 9 in the side room of a restaurant. The boys were familiar with philosophising and acquainted with the kind of questions. Their parents had been informed and asked for their consent; participation in the special format of the tetralogue was scheduled out of school hours. After the first tetralogue, the boys were asked for their bio data, their GPA language, completion of the Raven test and NEO FFI questionnaire, and to fill out the four evaluative post discussion questions. This questioning and testing procedure was not repeated later on. The results of the first two tetralogues were processed in previous chapters also.

Table 8.a. Overview follow-up tetralogues.

Date	Chron. (weeks)	Topic	Cat. topic	N part.	Duration (min.)	PQ
24-05-00	1 (1)	Live forever in perfect deep-frozen position	2	4	46	285
06-07-00	2 (7)	Written report and real life experiencing	1	4	45	260
17-10-00	3 (21)	Can a cat to become a champion?	4	3	32	267
					45*	373*
17-10-00	4 (22)	Is a champion necessarily the best?	4	4	23	171
					46*	341*
16-01-01	5 (35)	Australopithicus and Pithecanthropus	2	4	46	244
08-05-01	6 (51)	Can birds be guilty?	3	2	44	212
11-05-01	7 (52)	Can something exist that is not observable?	1	3	45	379
24-12-01	8 (84)	The real Ferrari	1	4	47	408
11-09-02	9 (121)	Free in prison?	2	4	44	280

Legend: Chron.: chronological rank-order (approximate intervals in weeks)
Cat. topic: number of category of philosophical topic (see Chapter 5)
N part.: amount of participants in tetralogue
Duration: duration of tetralogue in minutes
PQ: philosophical quality of the group (tetralogue) performance
* adjusted for duration

Description of the succession of tetralogues

The *first tetralogue* at Week 1, dealt with a question of everlasting life: is it possible to live forever if a perfect deep-freeze technique is available to preserve your living body? The tetralogue was loaded with many different anthropological, epistemological and metaphysical themes. A substantial number of authentic thoughts, original analogies, funny reasoning and category formation of living entities were mixed up. For example, the time taken to deep-freeze a human was compared to the time taken to freeze a carton of spinach. If the latter carries an expiry date, then the first must also. The period of being frozen was imagined as a time of being half alive and half dead. Because of the concept of 'half,' it was

argued that one might deduce the age of a frozen person as half the duration of being frozen. Finally, the boys wondered about the age of mummies and fossils as calendars differ from different points of view.

The *second tetralogue* at Week 7 took place at the close of the school year. The end of the last school year for the boys coincided with the end of being together for eight years in the familiar, trustworthy setting of the primary school. The initial question focused on the difference between a written report of a natural disaster in California and experiencing the same event. The epistemological topic jumped back and forth from the description of prehistoric circumstances to time machines, from futuristic images to recollections of non-verifiable events, from dreams to commercials, from imitated knights and their castles to photographic pictures. The venue of the tetralogue was moved because the school had to clean the original classroom before the holidays. This new classroom was only half cleared and the remainder quite a mess. The tetralogue was also interrupted by an outsider looking for something. The intimate atmosphere and the unusual classroom setting provoked many anecdotes tumbling over one another. Trains of thought concerning knowing, truth, and objectivity were created. At the end no convergence at all appeared as the boys could not keep their minds on one subject and easily skipped from one to another.

The *third tetralogue* at Week 21 was scheduled for October the same year after the boys had entered new schools. It was located in the primary school and scheduled after school hours. The boys were very anxious to meet each other as they were charged with fresh experiences. Unfortunately, one of the boys (Steve) did not appear. After waiting for a while, we decided to start the tetralogue with three participants. The initial question was: Is it possible for a cat to become a champion at catching mice and can such a cat get into the Guinness Book of Records? In the discussion, the boys searched for the meaning of *winning*: is *winning* always a honestly performed occurrence? Can *winning* happen by chance? What is the ultimate evidence of *winning*? Is the competition open to all? What about comparison between humans and animals, or between capable and disabled people? Is the best really the best? After 32 minutes, the fourth boy arrived. We finished the tetralogue session with three boys and started another that same afternoon, although scheduled as Week 22.

The *fourth tetralogue* at Week 22 with all four boys was initiated with a fresh question following naturally from the line of thoughts of third: is a champion necessary always the best given that the possibility remains that somebody exists who performs better? The attribution of being the best may be the result of bad or good luck, or of an accident. Attribution of being the best is impossible if the result cannot be communicated. What means anyhow being qualified as the best? Both tetralogues 3 and 4, contain many original lines of thoughts. Things that go without saying turn out to be less evident than initially seem to be and the reasonableness of final conclusions seem to be, of arbitrary nature, or based on artificial arrangements. Unfortunately, the utterances of one of the boys (Erny), were denounced by the others, especially after entrance of the fourth boy (Steve). This session took only 23 minutes because the boys had prior appointments after that time.

The *fifth tetralogue* at Week 35 focussed on the difference between Australopithicus and Pithecanthropus, with the question: 'Is the fossil link between ape and mankind of animal or human nature?' The discussion was more or less ruined by remarks of the boy (Steve) who neglected the third tetralogue. He made many dubious remarks and ridiculed almost all utterances, especially those of the boy (Erny) who was more or less excluded from Tetralogue 4.

Steve also did not appear in the *sixth tetralogue* at Week 51. Moreover, another boy (Tom) seemed to have forgotten the appointment. At this time it was not certain whether the series of tetralogues could be completed. Regardless, I decided run the tetralogue with the two boys who philosophised about the possibility whether birds could be guilty and held accountable for transferring fowl pest. The tetralogue included a fine example of analogy reasoning between men and puppets or robots. The boys imagined that if there is a God pulling the strings and running the show, then humans could impossibly be guilty. Robots cannot be guilty because they are programmed and controlled by humans. The question with respect to birds remained open. It was not possible to assess the reliability of the tetralogue (duologue) indicator scores because it deviated with only two participants (see Chapter 4, Section 4.3). A PQ index was not calculated because this calculation required the mean pq index of three participants.

In spite of frequent communications, a clear appointment and a promise, the boy who was absent in tetralogues 3 and 6 did not show up at the *seventh tetralogue* at Week 52. Making appointments with the four boys was becoming more and more difficult because of their different time tables and engagements in different environments linked to different schools and leisure activities. The initial question was in response to the fairytale about the new clothes of the king: Can things exist that are not observable? After exploring what can be observed and what cannot (air, transparent things, Martians), and imagining an invisible football that exists because it is moving air, it was agreed that existence appears to depend on knowing and believing. The boys tried to distinguish belief and knowledge and tested several examples. Experiential and rational positions relieve each other. Finally, they decided that existence is what is agreed upon by people.

The *eighth tetralogue* at Week 84, took place in the house of the author with all four boys because the primary school building was no longer available in free hours. The initial story was about an old Ferrari which subsequently had all its parts replaced. Since new parts were used to repair the car and old parts were recollected to reconstruct a car, the question was which Ferrari was the original one? The boys were very motivated when discussing this topic and elaborated on the ideas of one another, changing their own opinions frequently. All problems raised were translated into realistic examples of other kinds, often resulting in the use of analogical reasoning. Consequences of one part of the reasoning were conveyed to the other. All classic causes of Aristotle passed in review as the principle of the car is the design, the designer, the material, and its driving force. This repertory extended with the biography of the car and the value of the person who is judging.

The last and *ninth tetralogue* at Week 121 took place in a small room at a restaurant. The initial question was: Is it possible to be free in prison? Most remarkably in this discussion was the change of opinions. In the first part of the tetralogue three boys answered the question with *no*, one with *yes*. At the end of the tetralogue, the last mentionrd boy and only one other judged *no*, while the other two said it is possible to be free in prison. As in earlier tetralogues, Erny was still excluded and his utterances were treated aggressively or simply neglected.

Notwithstanding the described circumstances partly resulting from the primary to secondary school transition, nine tetralogues were performed with the same participants in almost the same composition (see Table 8.a). Because the selected boys were initially peers from the primary school, an atmosphere of familiarity was preserved during this follow-up period, probably in contrast to similar discussions with present-day classmates. A goal of the project was to detect changes in philosophically qualified thinking patterns: in pq and PQ indices. Based on previous observations, these indices are expected to decrease. In addition, the focus was directed on the course of indicator frequencies within one participant performances over time to look for intra-individual stability.

4 Explorations of the longitudinal data

Standardisation of PQ and pq indices

A separate SPSS file was created consisting of all variables concerning tetralogue performances of the four boys during nine sessions, completed with the time distance between the tetralogues calculated in weeks (approximately).

As can be seen in Table 8.a, not all successive tetralogues have an equal number of participants, nor an equal duration. In order to make longitudinal comparison of PQ indices possible, these values need to be put on the same scale (i.e., they need to be standardised). PQ of the group performance of Tetralogue 6 with two participants was calculated through use of the mean pq index of these two boys. Because tetralogues 3 and 4 took almost half the time of other tetralogues, raw values representing accumulated frequencies of utterances and indicator performances in the duration time of a tetralogue are not directly comparable with those of full tetralogues (45 minutes). For this reason, results of tetralogues 3 and 4 are converted to a standard duration time of some 45 minutes: values reflecting indicator frequencies performed during tetralogue 3 (duration time 32 minutes) and tetralogue 4 (duration time 23 minutes) are multiplied respectively with 1.4 and 2. PQ indices of group performance of tetralogue 3 and 4 are adjusted as well with respect to accumulated frequencies of qualified utterances and dialogical events.

Because pq indices of individual performances are based on the mean performance of single utterances, adjustment of them is not necessary.

Longitudinal results per participant

Three types of results will be shown: 1) changes in individual's pq indices over 28 months; 2) changes in PQ index (group performance) over the same period; and finally 3) changes in individually performed indicator frequencies.

Figure 8.1 shows four graphs representing changes in the pq index of each participant, and one graph representing the change in PQ index of the group performance over a period of 120 weeks. As PQ indices run from 200 to 500, these figures are divided by 100 to fit in the graph.

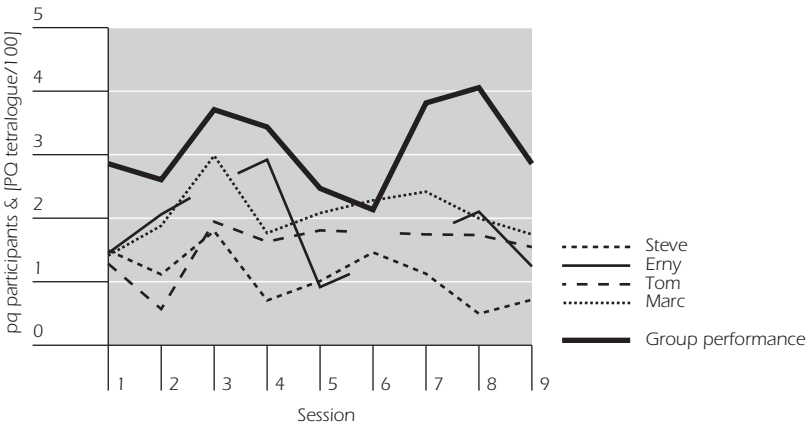

Figure 8.1. Changes in pq indices per participant and in PQ index of group performance.

Changes in pq indices over time

In analysing pq indices of each participant separately, no significant (linear) correlations were found between the individual philosophical quality and time (in weeks). Since the number of measurements is rather small, statistical *significance* may not be a good basis in this exploratory study for making statements about trends. When the correlation coefficient is conceived as a measure of fit between the observed data and a trend-model for the data, then the size of this correlation coefficient can guide in selecting an appropriate model. In spite of a general absence of significant linear correlations, the lines representing the boys' pq indices show some isomorphism. Trends may become visible by looking through linear, quadratic and cubic models representing the course of pq indices over time. The adequacy or fit of a model is indicated by the value of R^2 (muliple correlation coefficient squared). Specifically, R^2 is the percentage of variance in the data accounted for by the model. Table 8.b shows the values of R^2 for each model for each participant.

For Marc, Erny, and Tom, a cubic trend seems to be a reasonable representation. For Steve, none of the models seems to be adequate. Also, the philosophical quality of the group performance (PQ) seems not to be well represented by any of these three models. Individual cubic trend lines are shown in figures 8.2a to 8.2c for Marc, Erny and Tom.

Table 8.b. Fit (R^2) of three models for the course of pq indices and of group performance (PQ) over a two year period.

Participants	Number of performances	R^2, estimated according to different trend models		
		Linear trend model	Quadratic trend model	Cubic trend model
Marc	9	0.00	0.33	0.44
Erny	9	0.32	0.32	0.35
Tom	8	0.11	0.43	0.50
Steve	6	0.06	0.11	0.12
Tetralogue (PQ)	9	0.02	0.07	0.14

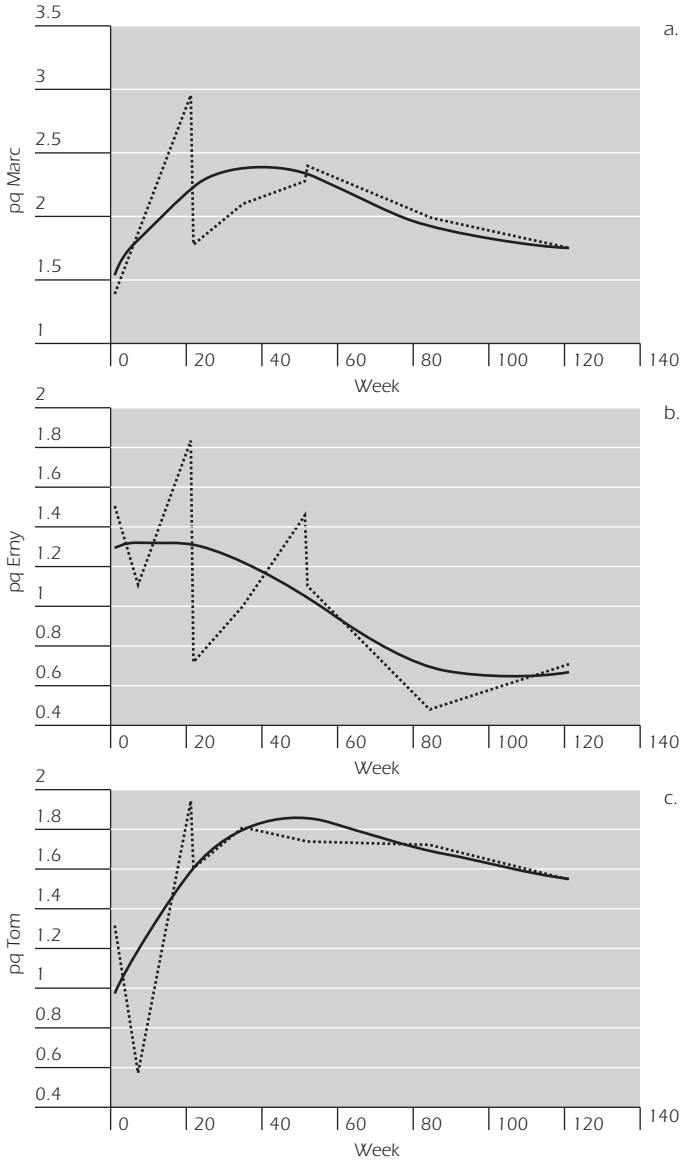

Figure 8.2a, b, c. Cubic estimations for the course of individual pq indices.

Individually performed pq indices, estimated by a cubic model show similar outlines for the three boys (figures 8.2a to 8.2c). These lines increase until somewhere between Week 20 and Week 40, then decreases from that time (i.e., briefly after the school transition). A decrease in pq indices becomes manifest in the boys after some time. This decrease can be substantial as Erny demonstrates. In primary school, Erny showed an estimated pq index of 1.3; after two years, his pq index was estimated to be about 0.6. The decline of 0.7 corresponds with an effect size (Hedge's g) of [0.7 / 0.5] = 1.4 (the pooled standard deviation for pq in Erny's age ranges in the VMBO group is 0.5).

A similarity becomes manifest when comparing the trends of individual pq indices in this follow-up study with the trend in the cross-sectional data of *all* participants with a low educational level (primary school and VMBO) in the thesis. The fit of a linear model for these cross-sectional data (N = 66, boys and girls) is very poor (R^2 = 0.02; p = 0.21), while a quadratic model fits reasonably better (R^2 =0.12; p = 0.02). The addition of a cubic term does not improve model fit. Figure 8.3 shows cross-sectional trend estimated according to a quadratic model, demonstrating an initial increase followed after about 12 years of age by a decrease in individual philosophical quality.

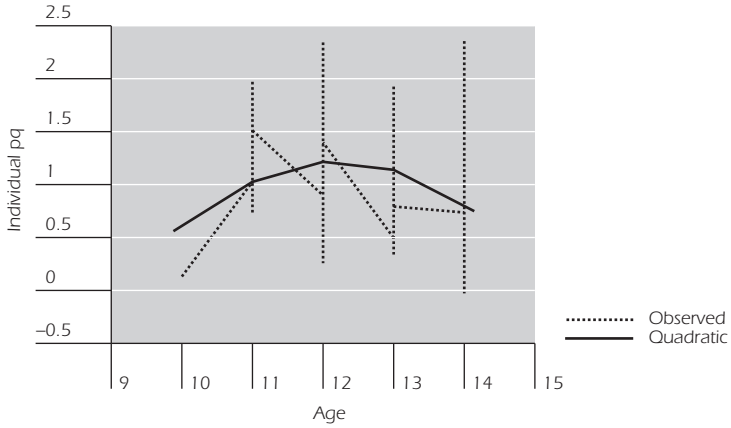

Figure 8.3. The course of individually performed pq indices of all participants with a low educational level (N = 66), estimated according to a quadratic model. Cross-sectional data.

Longitudinal changes in PQ indices over time

No systematic changes could be detected in the course of philosophical quality of the nine group performances (PQ indices). Table 8.b (last line) shows that none of the three models fitted well.

Changes in indicator frequencies over time

Figures 8.4 to 8.7 show the progress of performed indicator frequencies for each participant during this follow-up period. Indicators admitted are: indecisive thinking

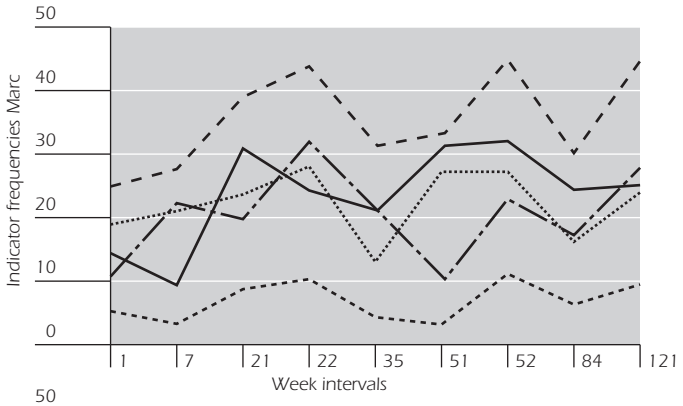

Figure 8.4.
Indicator frequencies in function of week intervals with respect to participant Marc.

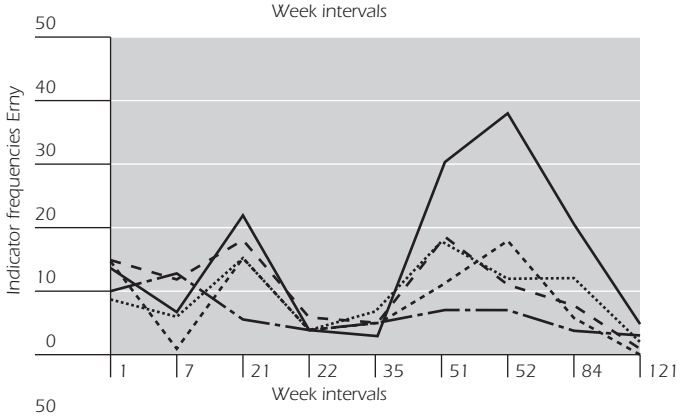

Figure 8.5.
Indicator frequencies in function of week intervals with respect to participant Erny.

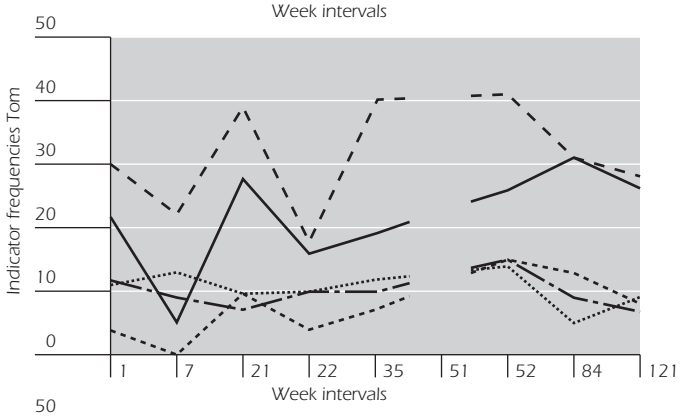

Figure 8.6.
Indicator frequencies in function of week intervals with respect to participant Tom.

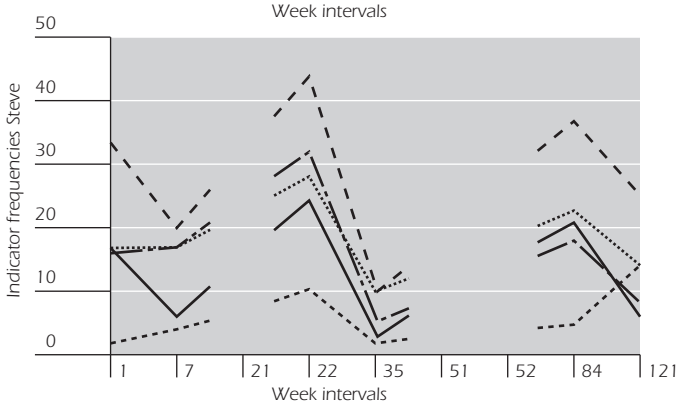

Figure 8.7.
Indicator frequencies in function of week intervals with respect to participant Steve.

(Idt), openness (Op), tentative behaviour (Te), epistemological position (Ep), and reasoning quality (Re). Frequencies of anecdotal quality are not reported here; anecdotal quality is only taken into account for results on the level of group performance.

Tentative behaviour is the most frequently scored indicator in this period, at least for three participants, and openness the lowest. These graphs, figures 8.4 to 8.7, demonstrate isomorphism of trends in the use of philosophical indicators. Per participant, the combined play of five indicators behaves similarly under changing circumstances. As was demonstrated in Chapter 5 (Section 5.3), indicator frequencies are not influenced significantly by different philosophical topics. Indicator frequencies show co-variation: an increase in one indicator goes together with an increase in the other. These outcomes are consistent with the statement that the five indicators refer a common underlying entity: individual philosophical quality (Chapter 4), indexed by pq.

Anecdotal quality

Apart from the five indicators mentioned, anecdotal quality is of interest because a decrease was expected according to the theory introduced in Section 8.1. A significant decrease in scored frequencies of anecdotal quality is detected in the performances of all four boys. A linear model captures the negative relationship between frequencies of anecdotal quality and time/age very well (r is, respectively for Steve, Erny, Tom, and Marc: –0,53, –0,59, –0,46, –0,50).

5 Conclusion and discussion

No linear change in philosophically qualified thinking patterns over time could be demonstrated with respect to individual performances, group performance and relevant indicator frequencies. However, when depicting changes in pq indices, more complex trends were observed. These trends are best described, in three of the four participants, according to a cubic model. They show congruence with respect to highs and lows: initially, pq indices increase during the last months on primary school and the first months after having left this school. Between three and five months after having entered a secondary school of low educational level, pq indices decrease. This moment is nearly the same in three of the boys. Inspection of cross-sectional data for VMBO-level pupils also showed a decrease starting after the age of 12.

Results in the light of expectations

These trends: a) do not contradict and corroborate my previous observations; b) are in line with some theoretical considerations, but not with all; and c) mirror environmental circumstances. According to my initial observations, the philosophical quality do not increase after entering a secondary school of low educational level

notwithstanding a rise in applying logical rules through internal reflection (Inhelder & Piaget, 1955; 1958). A tendency to thinking and speaking less in terms of concrete situations is reflected through a significant decrease of anecdotal quality scored in utterances during the follow-up study. This decrease in references to real life experience is concordant with the increase of formal thinking patterns at the expense of representing concrete things and events as objects of thought. Although naïve developmental models stress rising capacities to abstract thinking, meta-cognition and academic knowledge, no increase of pq indices is found. Obviously, the observed changes in philosophical quality deviate from the progress of mentioned cognitive qualities. This outcome confirms the independent identity of a philosophical quality examined in chapters 4 and 6. However, the decrease in pq indices may also indicate a regression in cognitive development. This has often has been observed before significant cognitive structural changes, according to Neo-Piagetians (Verhofstadt-Denève, Van Geert, & Vyt, 1995).

Environmental circumstances

At primary school, differences in average philosophical quality (pq) were not detected between pupils labelled for future high and low educational level. At secondary school, a difference is found between pupils of high and low level schools, implying some influence of the school environment. The results of this follow-up study show trends in the performance of philosophically qualified thinking patterns over time. From age 11 on, an increase in formal thinking patterns is assumed on theoretical grounds (Inhelder & Piaget, 1955; 1958). In the case of entering a VMBO school (after the age of 12), a decrease of academic performance is assumed to be the result of environmental circumstances. This is emphasised by Vygotsky (1964), in the theory of affordances (Jackson, 1995), and by the stressful transition experiences, especially in case of poor achieving youngsters (Berk, 1997). The results obtained (decreasing trend starting at age 12) are consistent with the hypothesis of the downgrading influence of the VMBO environment on individual performed philosophical quality. More rigourous evidence of the influence of educational environments can only be obtained by better designed and experimental studies with control groups (to control for e.g. 'regression to the mean').

However, qualitative observational descriptions of tetralogues reveal some remarkable environmental circumstances, for example: disrupting behaviour, neglecting appointments, and being denounced. These events cannot be described as random accidents and may very well clarify changes in individual performances over time. The denouncement of Erny, after transition from primary to secondary school, seems to be reflected in his very low pq index (see Figure 8.1), beginning in the fourth tetralogue. From then on, this boy shows noticeably fewer utterances than his peer-participants and a downward trend is visible in the course of his pq indices (Figure 8.1). Steve's disrupting behaviour accompanies low individual pq indices of his peer participants. At the same time, he seems to induce relatively high qualified group performances (PQ indices) by scoring the highest individual pq index. Tetralogue performances 3, 6 and 7 during Steve's absence show relatively high pq indices by the other three boys. Some months after having left primary school,

a turning point can be indicated in three perspectives: statistically through downward trends; biographically through school transition; and circumstantially through the described events in the environment. Individual performances may be described very well as reacting to the environment, especially to the performances of peer-participants. This was been emphasised in chapters 5 and 6 with respect to group composition. Here, the shared environment is stressed.

Fluctuations in philosophical quality

The observed fluctuations in individual philosophical quality are probably related to habitual environmental circumstances. In Chapter 1, psychological research was cited in which changes in talents were examined in close relation to changing contextual support, changing constraints, changing tasks, and facilitated by changes in the individual's environment. Also, some environments may impede talent. The results of this follow-up study do not contradict the concept of philosophical quality in terms of such talent. Further exploration is needed to clarify changes in qualified thinking patterns after entering secondary schools of low educational level.

9 Philosophical quality in perspective

In this chapter, the results of this study are evaluated with respect to research aims and expectations. Limitations that may influence the conclusions will also be reviewed. Proposed indices for philosophical quality will be discussed in the context of selected scientific traditions and current discourses. These include a philosophical tradition, ideas about rationality, and studies concerning philosophising with children. The contributions made by this research to the study of these domains will be described. Referring to Russell's labelling of philosophy 'the no-man's-land between science and theology', non-rational but philosophical qualities will be taken into account. The results of this thesis will be evaluated against empirical psychological approaches to wisdom, cognitive development and moral development. The appropriate use of *talent* as attribute to philosophical quality will be discussed. Finally, some proposals are made for the use of measures of philosophical quality.

1 Philosophical quality, pq and PQ

Assessment of philosophical quality

This thesis has identified a specific quality of thinking patterns, namely: philosophical quality. Initially, three main features of philosophy are derived from a review of history and the works of well-known modern philosophers. The thinking patterns of youngsters that deal with philosophical questions were observed and matched to the main features of philosophy to construct a conceptual framework. Five convergent indicators are developed that can be observed and counted in the oral expression of thinking patterns. Performed thinking patterns are investigated through the use of *tetralogues*: i.e., standardised situations of philosophical discussions. The utterances of youngsters (aged 11 to 18 years) during tetralogues are scored on the presence of indicators. Objectivity and reliability of the scoring method is demonstrated by inter-rater agreement, internal consistency of indicator frequencies, and by good split-half and test-retest coefficients. By concatenation of indicators in line with the main features of philosophy, two numerical indices for philosophical quality are created: pq and PQ indices, reflecting the philosophical quality of, respectively individually and collectively performed thinking patterns. The construct validity of these measurements is corroborated by a significant correlation between calculated indices for group performed philosophical quality and previous estimations of tetralogue's philosophical quality ($r = 0.54^{**}$), and by replication of nomologically expected relations as exposed in Chapter 6 (see figure 6.3 and 6.4). So, philosophical quality performed individually or in groups is expressed in quantitative terms.

Mental endowment

The project also explores philosophical quality in relation to participant characteristics and identifies a link between this quality and its performer. Philosophical quality is related significantly to *openness to experience* as personality trait ($r = 0.27^{**}$); to educational level ($r = 0.44^{**}$, with respect to individual performances); while a moderate relationship is found between philosophical quality and intelligence ($r = 0.15^{*}$) indicating that philosophical quality diverges considerably from convergent thinking. Average philosophical quality in grammar schools is significantly higher than that in schools for VMBO. At the same time, no relation between philosophical quality and age is found. There are indications (see chapter 6 and 8) that individual philosophical quality may be related to habitual environmental circumstances that may hone, feed or discourage this quality. The presence of this quality already in childhood, and of stable individual differences, its independency of age, its relation to openness to experience, and its readiness to be exploited in cooperation with the environment, are characteristics of individual philosophical quality likely to be a mental endowment.

Limits to generalisation

The extent to which the performances of youngsters participating voluntarily reflect those of all youngsters can be questioned. Can the results be generalised and applied to non-participating youngsters in the participating schools, or even beyond The Netherlands and Belgium? Selection criteria are described in Chapter 4, covering a broad spectrum of participant characteristics with one shared feature: participants are volunteers. Participation in tetralogues was generally high: for a class with an average of 25 pupils, between 0 and 10 percent of the students did not volunteer. Some students may be not have been sensitive to the thinking patterns required to participate in the tetralogues; others clearly come from exotic cultures. However, the purpose of this project was to demonstrate that philosophical quality can be measured and not to make general statements about the thinking patterns of potential performers. Nevertheless, a few issues deserve attention: language, culture and characteristics that deal with multicultural group composition.

Language and culture

Pupils of allochtonous origin probably did not have a thorough command of the Dutch language: a condition sine qua non to participate. If the list of expressions (Appendix II) to be scored is adjusted to language, it is reasonable to think of philosophical quality without regard to native language. This does not hold for culturally influenced attitudes that make pupils ill-disposed to a serious and open approach to open-ended questions. It must be admitted that most non-volunteers (non-participants) were allochtonous pupils even though they understood and spoke the Dutch language. Few volunteered and agreed with the tetralogue conditions. Observation of philosophical practices in primary schools, made by the author over many years, reveal that many Moslem boys are short on openness and fortify

themselves in expressions of clear-cut contents. Here, another hypothesis formulated earlier is demonstrated: conformist attitude and frequent confrontation with authoritarianism. Allochtonous families (Chinese, Turkish, Moroccan, Creole-Surinamese) in The Netherlands show more authoritarian supervision in their pedagogical style, while an authoritative style is becoming more and more adopted (Pels, 2000, p. 207). One tetralogue (out of 95) begins with a Moslem boy expressing an authoritarian attitude and intimidating a Moslem girl. The Moslem boy shows the lowest pq index (0.70) in this tetralogue, while producing the greatest number of utterances (98). Mean pq index in his category of primary school pupils: 1.20 (s^2 = 0.33; range from 0.12 to 2.43; N = 53). At the same time, the girl scored pq index: 0.74 in 38 utterances. This outcome is comparable to tetralogue results where participants demonstrated some degree of authoritarianism. Authoritarianism and conformist attitudes seem relevant with respect to the performance of philosophically qualified thinking patterns, and need to be explored in future research.

Characteristics dealing with multicultural group composition

This project was restricted to different geographical locations in The Netherlands and Belgium. It is conceivable that philosophical quality is a residual product of sophisticated urban ways of cognitive behaviour. Activities of philosophy usually occur in cities rich in mental challenges and fertile of multicultural contacts (Verhoeven, 1973). In this belief, multicultural group composition may positively influence the philosophical quality. For this reason, tetralogues aimed at multiformity of life courses and gender composition. Unfortunately, a relationship between such group composition and philosophical quality of group performance could not be identified because of the nebulous assessment of regular life course as a participant characteristic. Multiformity is also at stake where tetralogues are composed of participants from different cultures. Living between different cultures, for example in cities or in rural areas, or away from exotic cultures may result in differences between performed PQ indices. To investigate possible divergences with respect to different geographic characteristics, different regions were initially selected to participate in this project. However, in The Netherlands and Belgium, participants appeared to be familiar with cultural differences. Because a relationship between philosophical quality and geographical background in the dataset of this study was not expected, this feature was not considered further. However, it is conceivable that other results may be found in areas of distinct or homogenous cultural composition. In brief, the outcomes of this research cannot be extended to adolescents from areas lacking cultural differences.

Philosophical quality of non-philosophical discussion

Naively, one may assume that indices for philosophical quality are only meaningful if they can discriminate philosophical from non-philosophical discussions. However, it is the claim of this project that philosophical quality has to be discriminated, but not the philosophical discussion. To check this, discussions were recorded in an observation study with pupils of 14 to 16 years old. They were asked to discuss

a given social dilemma relevant to their age situation. Because these discussions were not standardised in the same way as in tetralogue format, and because there were many leaps in topics induced by an outsider, pq and PQ indices cannot be assessed and results cannot be compared with tetralogue results. Notwithstanding this drawback, indicator frequencies were scored for two discussions with four participants. Adjusted for duration time, mean frequencies of indecisive thinking, openness, tentative behaviour, epistemic position, and reasoning quality come up to respectively 29, 15, 72, 115, and 42; whereas figures assigned to tetralogue performances amount to 43, 28, 80, 58, 48. So, except for epistemic position, lower frequencies are measured in case of non-philosophical discussions. Although a non-philosophical occasion may negatively influence indicator occurrences, the difference between philosophical and non-philosophical discussions is not a difference of *kind* but one of *degree*.

2 Philosophical quality in the perspective of philosophy

This study has identified a philosophical quality in thinking patterns. Philosophy emerges as conceptual clarification and epistemological comments on questionable truth claims (Philipse, 2004). Philosophising has been analysed through its historical mainstream, through the work of modern philosophers, and through close observation of adolescent thinking patterns in response to philosophical questions. These analyses have led into the detection of five indicators and the construction of indices for philosophical quality of individual and of group performances. Since these indices represent philosophical qualities, characteristics of philosophical thinking patterns can be clarified and specified. This experimental study confirms many of the cherished generalisations of philosophy by philosophers. It also corroborates Hirst's speculative statement of philosophy as a specific different knowledge domain (1980). No other attempt to quantify *philosophical quality* could be traced in academic philosophy, probably because of an apparent paradox dealing with supposedly intangible philosophical qualities.

Quantification of philosophy and its perceived intangibility, a paradox

Assessing philosophical quality, particularly in quantitative terms, at first sight seems to contrast with the commonly perceived characteristic of philosophy by its lack of a clear, intangible and final goal. This apparent discrepancy has been discussed in Chapter 3, Section 3.4. It is also noticed that a measurement is admitted to describe philosophy, to evaluate philosophical essays or discussions, and to remove philosophy from interminable thinking patterns. The contours of philosophy must be determined and its content expressed in quantifiable terms. By developing a measurement tool, this study attempts to contribute to this goal. Here, philosophical quality is measured through the quantity of occurrences of sensitivity to vagueness, ambiguity, and uncertainty, openness to the unknown, searching for oppositions and borderline cases, while stripping concepts of their rigid meanings. The empirically achieved results reflect manageable approximations to the *philosophical quality* concept, without the presumption of having covered the entire concept.

Two more questions must be discussed with regard to measuring philosophical quality against a background of philosophy in general: its relation to common sense ideas and its relation to existential life situations. In this study, philosophically qualified thinking patterns are not traced according to academic norms that stress convergent thinking. They are also not traced according to norms that presuppose philosophy is exclusively connected with thinking about life, death and god. It is noticeable that many youngsters are initially inspired by philosophy in this way. Moreover, the collected tetralogue data are open to score expressions containing concepts of *life and death* and *god*. It may be of some relevance to relate these occurrences with pq and PQ indices to establish the relation between the different notions of philosophy. In addition, it would be interesting as well to trace potential links between such occurrences and participant characteristics. This should be object of further investigation.

Existential life situations

Philosophical quality of the content of discussions sometimes stands out against an existential background. It was hypothesised that imprisoned, adopted and disabled youngsters struggle with their delicate life situation, and would consequently demonstrate a substantially higher philosophical quality. This expectation is not supported by the data, possibly because of poor assessment of the life course of 'regular' youngsters. Nevertheless, philosophical discussions with 'irregular' participants have noticeably unique thinking patterns. This can be illustrated through discussions about the relevance of communication in relation to life with physically disabled youngsters. For example, a severely physically disabled 19-year old boy (*spastic tetraparesis*) indicated that his handicap was part of his identity. Therefore, he reasoned, his difficulty to communicate and contribute to society was not a limitation. He declared that, even with a choice, his handicapped existence, including his continuous dependence on wheelchair and on many helpers and aides, was the best form of life imaginable. Existential qualities are also shown in discussions with youngsters imprisoned in Teylingereind. Here, it was most striking the extent to which predominantly Moslem boys were eager to discuss (openly) the authority and power of god, yet were disgusted when discussing conscience and inner 'voices'. In both instances, initial philosophical questions are conditioned by the absence of definite answers. The thinking patterns performed touch existential questions but may be identified without philosophical qualification. The ease of attributing philosophical quality of these discussions may be credited to biases and prejudices of the outsider observers or a discrepancy of presuppositions between participant and observer. However, the inferred philosophical quality of these contents is not born out by the results of pq and PQ indices.

3 Philosophical quality related to rationality

In terms of rational thinking, philosophically qualified thinking patterns refers mainly to reasoned divergent thinking; but also in a small part to convergent thinking or academic skills. Philosophical quality has a broader scope than mere

rationality, as philosophical indicators also identify attitudinal aspects of non-rational origin. Non-rational thinking patterns covered by philosophical quality are still expressed through the rational means of language performance, although pq and PQ measurements do not measure the content of its propositions. Philosophical quality (pq and PQ) shows family resemblances with rationality and encloses some non-rational elements.

Because the use of *rationality* is ambiguous in different research domains dealing with thinking patterns, an explanation is needed here for this concept. In a narrow sense, rationality is a characteristic of thinking patterns, generating beliefs on the basis of appropriate reasons. It excludes judgements and reasoning leading to contradictions (Dawes, 2001). As in a classical conception, rationality, must comply with formal logic and mathematics yielding universal and necessary results. The problem with this view in the context of this project is that it would exclude rational disagreement whenever the same piece of evidence generated incompatible conclusions. Moreover, it could be questioned whether all matters of value are subject to rationality without being paralysed by considerations of relativity. As stressed above, not all thinking patterns need to converge. Divergent thinking, such as creativity, openness to the unknown, and sensitivity to ambiguity may emerge from a common need for orientation in an everyday world based on common biological roots, common ways of perceiving, and common use of if-then-relationships. Consequently, if rationality is conceived according to such a broader view, all *attempts* of reasoning, according to formal or informal logics, or through analogies, all types of wonderment, sensitivity to ambiguity, vagueness and uncertainty, and tentative behaviour in the context of philosophical inquiry, may be labelled as rational.

Non-rationality

Seemingly non-rational approaches to philosophical questions remain open to rational evaluations and are often testable. How do we label the exclusive quality of non-rational approaches of reality? What does it look like? One may refer to the imaginative thinking patterns in Chapter 3 where a machine is staged that can defy gravity and a formula is imagined that could make magnified objects really large. One may imagine types of emotional 'considerations' for dealing with reality and acting upon experience. One may also stress the quality of arguing through emotion, compulsion and unintelligible beliefs. According to Neo-Piagetian views, rationality is underlying many of these non-rational approaches. The intelligibility of expressed thinking patterns and of observable behaviour in general makes such non-rational approaches open to rational judgement.

Thinking and expressing

Another fundamental problem concerning the roots of thinking patterns deals with the relation between thinking and its observable expression. Philosophical quality in this study is a discursive quality because it is uttered through intelligible

expressions. We can never be sure that thoughts expressed really correspond to semantic thinking or that the individual is fair in their intention to express these thoughts. This gap between thoughts and utterances should always be kept in mind when valuing observed human behaviour in general. Furthermore, an ambiguity emerges if somebody is saying nothing. In this study, expressions without any indicator occurrence are excluded from philosophical qualification. A philosophical index of *0* does not mean a zero potential philosophical quality or a non-thinking quality of its achiever, but rather a *0* score is inaccessible to measurement. To cope with some of the limitations in verbalising, a conditional requirement was introduced and some non-verbal meanings were recorded: for instance, intonation and body language. One of the selection criteria for participation in tetralogues was the ability of youngsters to understand and verbalise their thoughts in Dutch.

Analogous to the judgement of non-expressed thoughts is the phenomenon of not knowing. Not knowing may be the assessment of knowledge absence or incorrect knowledge. However, indices for philosophical quality, especially indicated through frequencies of indecisive thinking, detect types of not knowing as inferentially processed outcomes of thinking patterns in response to unanswerable questions that often stress the arbitrariness of a perceived reality.

Rationalism

Rationality must be distinguished sharply from philosophical rationalism. Philosophical rationalism is opposed to empiricism and refers to an epistemic stance explaining knowledge through the mind (cf. Descartes) or through experience (cf. Locke, Berkeley, Humes). It is relevant to stress this point here because the content of many tetralogues touches on the difference between these epistemic stances. It is noticeable that five tetralogues in this project show a clear competition between a rationalistic, and an experiential approach to reality and knowledge. In these tetralogues, invariably a debate develops between two girls representing the experiential attitude, and two boys with a rationalistic stance. For example, the girls argue against the existence of time in case of a reality without human beings because nobody can experience it; the boys argue in favour of the existence of time independent of human presence. In another tetralogue dealing with quality of art, the girls argue in favour of judging artistic value according to the experience of producer and observer; while the boys stress the quality of the product independent of the experience. A third tetralogue debated the existence of paranormal phenomena. In the boys' view, they cannot be explained according to rationalism. Checking for separate indicator frequencies and for pq indices, no significant differences between the two positions can be noticed. Philosophical rationalism and philosophical empiricism are expressed to similar degrees in the results. It may thus be concluded that philosophical quality is independent of rationalism as a philosophical stance.

4 *Philosophical quality vis-à-vis philosophy with children*

The results of observations and records of thinking patterns expressed by relatively young people is evaluated in the context of twenty years of experience in philosophy with children. Some countries now have well-established philosophy programmes, but only for primary school children. In this study, out of the total of 215 youngsters with computed pq indices, 53 are from primary schools. An additional 87 pupils, from 17 to 20 years in age, participated in tetralogues that have yet to be analysed. To date, quantitative research data have not been published concerning philosophically qualified thinking patterns in secondary school pupils.

Philosophy for Children

The programme *Philosophy for Children* has been running in primary schools for more than thirty years. It started as an experiment in a small commuter town in the United States and developed into a worldwide movement (Lipman, 1976). Since then, a number of significant publications has been devoted to this subject, including research materials, books and journals promoting and developing the programme. Most texts are of a prescriptive or descriptive nature. Very little research has directed at studying effects of the programme on children. Neither has there been a rigorous analysis of the quality of philosophical thinking patterns in youngsters outside the programme. Only a few researchers have noticed the relevance of this quality.

Effects of the Philosophy for Children Programme

When evaluating the effects of 'philosophising with children', external and internal effects should be distinguished. Internal effects refer to progress in philosophical thinking. A pioneer attempt at measuring such progress was the *New Jersey Test of Reasoning Skills (NJTRS)* (Shipman 1983), concentrating on formal reasoning. Except for being limited in scope to identifying reasoning skills, the content of this test appears to be dependent of the curriculum and cannot determine to what degree children are able to reason given the aims of the programme. Incidentally, a correlation between data from *NJTRS* and academic grades in mathematics and language has been established (Cebas & Moriyon, 2003). External effects of philosophical discussions in the classroom focus on improvement of performance in other school subjects, like mathematics and foreign languages (Lipman 1976, Schleifer and Courtemanche 1995); on the influence on social attitude in the classroom (Pálsson 1995; Niklasson, Ohlsson and Ringborg 1995; Gazzard, 2000), and on developing a democratic attitude by liberation from stereotypes (Cwi 1976; Pálsson 1995; Schleifer and Poirier, 1995; Cam, 2000). However, such researches do not always clarify why such external effects are specific to philosophical discourse (Santi 1993, 1995; Yule and Glaser 1994). Moreover, these evaluation experiments have not demonstrated internally effects (Santi 1995, Cebas & Moriyon, 2003).

Considering the quality of philosophising children

Only a few researchers have considered the philosophical quality of discussions with children, stressing the philosophical content in a very informal manner. According to Santi (1995), qualities can only be observed when practicing philosophical discussion, providing these with comments. Pointing to the philosophical content of a discussion and commenting on the way children philosophise alone does not satisfy a measurement of philosophical quality of thinking patterns. Yule and Glaser (1994) analysed transcripts of philosophical discussions through a list of analytical skills designed by Lipman. Their focus was on the quality of the dialogue and they recognise the importance of the 'role of imagination, stories, anecdotes', and state that these categories have to be analysed in more detail. This study demonstrates that assessment of philosophical quality is possible by identifying indicators and calculating indices in a standardised testing procedure.

Judging the philosophical roots of children's thoughts

Most commentaries on 'philosophy with children' in literature address educational goals for improving skills, attitudes, and qualities that are potentials of salutary effects, while advocating philosophy in the classroom. Few observe and judge children without such bias or focus on philosophical topics. Matthews (1980, 1994, 1998) is one of these exceptions, analysing philosophically qualified thinking patterns of children in primary schools. Another example: Freese (1990). Freese describes thinking patterns of children derived from literature and from his own youth, although the reliability of his approach is questionable. Both authors consider children's thinking patterns on philosophical questions as a triangle between philosophy, child, and cognition without any goal orientation. Neither Matthews nor Freese provide a measurement tool for philosophical quality and their descriptions and analyses do not refer to any standardised testing procedure. Despite this lack of proper quantification, their results may be compared with those of this study.

Matthews' approach

Matthews begins by stressing the authenticity of children's thinking patterns recognised as philosophically qualified by professional philosophers. Next, he compares these thinking patterns with Piaget's stages of cognitive development. Authenticity and inventiveness do not fit the presupposition of evaluative judgements to which the concept of *childhood* is submitted. 'Childhood' acts as a specification of the philosophical domain and thus is comparable with other such specifications as *'life'* and *'men'* are specifications of philosophical anthropology. Therefore, Matthews argues in favour of 'philosophy of childhood' as an autonomous (sub)discipline (Matthews, 1994).

Mathews evaluates openly and without the requirement of correctness children's thinking patterns against potential cognitive development, especially according to Piaget. Contrary to his view, empirical studies suggest that children's thinking

patterns develop into adult patterns, presupposing a final (correct) stage where cognitive operations are mastered. Tracks of thinking are supposed to be right or wrong. However, the thinking patterns of children may develop beyond well-defined steps, and sometimes they are even successful by accident. Piaget does not evaluate these patterns by psychometric measurements. Consequently, a difference must be noticed between the philosopher's view and the developmental psychologist's view. The first views individual thinking patterns in relation to the development of concepts into alternatives within a philosophical tradition of thinking, in the manner of Deleuze and Guattary (1994). The second views thinking patterns in relation to the human development of correct thinking patterns that accord to adult norms. Philosophically speaking, definite assessment of correctness of thinking results is impossible. Philosophically qualified thinking patterns show some level of disquietude, of not knowing or at least of being conscious of the tentative and arbitrary nature of concepts. Do colours exist when nobody can see them? What would be the truth about the conservation of substance beyond our world? According to Piaget and Inhelder (1974), children explore the conservation of substance through ideas about coherence between weight and volume; these ideas show a stage-wise approximation to adult norms. From a philosophical point of view, the physical truth is open to alternative interpretations.

Thoughts and systematic trains of thoughts during a philosophical quest drive the philosophical object of this project and determine the values of the measured parameters, as in Matthews' analysis of children's thinking patterns. Neither development of concepts nor development of child is taken for granted. This research shares the qualitative approach with the analyses of Matthews and Freese and provides these with quantifiers of philosophical qualities to get a handle on philosophically qualified thinking patterns of children and adolescents.

5 *Philosophical quality in the perspective of empirical studies*

In this study philosophical quality has been investigated empirically and may consequently relate to concepts and constructs in other empirical fields.

Wisdom

In Chapter 2, the main features of philosophy and wisdom, conceived as an attitude towards knowledge, are related. In ancient philosophical tradition, wisdom refers to a life stance of inquiry and to consciousness of not knowing, aimed at collecting a broad spectrum of knowledge and experience that leads to 'arete'. Today, wisdom is defined as understanding uniting reflective attitude and practical concern (Kekes, 1995); expertise in life pragmatics (Baltes, 1990, 1995, 2000, 2004); balanced application of intelligence, creativity and knowledge (Sternberg, 2003, 2004); expertise in uncertainty (Brugman, 2000); emphasising features of problem identification (Arlin, 1975); reflective thinking (Kitchener, 1983); a stance halfway certainty and doubt (Meacham, 1983); and dialectical thinking (Riegel,

1973). Like wisdom, philosophical quality concerns the measurement of, and the way statements of thought are delivered to expressions. Philosophically qualified thinking patterns also emphasise divergent thinking, uncertainty, metacognition, and context dependency. However, they do not include the content of statements, affective components, and types of behaviour.

Both philosophical quality and wisdom can be measured. The measurement tool developed here for philosophical quality operates through indicator frequencies. Indicators like sensitivity to ambiguity, vagueness and uncertainty (Idt), openness (Op) and epistemic position (Ep), appear to identify metacognitive qualities of inquiry and not knowing. Moreover, pq and PQ share characteristics with wisdom measurements. According to Brugman (2000), wisdom is independent of age; is related with educational level, and is linked to *openness to experience* as personality trait. Overviewing the variety in wisdom research, different types of wisdom aspects are emphasised and reflected in the number of measurements. According to the Handbook of the Psychology of Aging, several types of wisdom assessment exist (Brugman, in press), based on a variety of approaches such as pragmatic (Baltes, 2001; Sternberg in progress), epistemic (Kitchener, 1993; Brugman 2000), and others (Ardelt, 1998; Happé, 1998; Webster 2003; Jason, 2001; Perry, 2002). It would be useful to investigate the relationship between these measures and pq indices to see to what degree they assess the same thing.

Cognitive development and education

Compared to regular cognitive educational practices, including morally loaded goals, philosophising with youngsters begins with questions that have no definitive answers and require openness to a multifaceted and unstructured reality. Identifying philosophical quality mainly involves tracing divergent thinking patterns that go beyond the convergent thinking as presented in intelligence tests. The relative independence of philosophical quality from intelligence is demonstrated by the low correlation between pq indices and Raven results. Thinking patterns that are not submitted to requirements of correctness and that express relations that do not fit in adult paradigms are no longer qualified as imperfect. Some may qualify these thinking patterns as romantic, pre-rational or naïve since they stress the non-serious and imaginative nature of many children's expressions (Elbers, 1993; see also examples in Chapter 3). However, pq indices deal with the qualities of divergence and imaginativeness without disqualifying them. Attempts to trace other characteristics of individuals that concern non-convergent mental patterns were undertaken in the 1970s with respect to creativity, but were not successful (Brugman & Dudink, 1976), and in the 1990s with respect to emotional and social capacities (Mayer & Salovey, 1990, 2000; Goleman, 1995; Sloan, 2000).

Moral development

Well-described attempts to measure cognitive qualities, stressing a moral content, have been undertaken in studies of moral development. These may be considered

relevant here since the discipline of ethics is included in philosophy. As demonstrated in Chapter 5 however, differences in philosophical quality are unrelated to differences between ethical and other philosophical topics. Individual pq indices represent philosophically qualified thinking patterns through indicator occurrences not reflecting any justification per se (Kohlberg, 1981; Olthof & Brugman, 1994), nor virtues. More or less, these approaches (Kohlberg, 1981; Olthof & Brugman, 1994) show convergence toward a goal orientation and, as a consequence, uncover developmental stages. They may only comply with the results of this research if tests for moral development detect unbiased qualities of cognitive capacities to reason about acts and thoughts.

Moral development involves more than judgement, reasoning, and justification expressed in propositions. Morally loaded expressions cannot be evaluated according to standards of meaningfulness as Wittgenstein (1969) stresses. Ethics is linked to direct, concrete acting, to an empathic attitude, and to some stoic remoteness. It shows itself in acting. 'Finally, actual moral behaviour is independent of the philosophically founded and scientifically justified self-image.' (Stokhof, 2001, p.32). Ethical discussions can also be approached according to their dimension of affective emotions (Stevenson, 1963). Stevenson stresses that moral judges express approval or disapproval and seek also to influence the feelings of approval and disapproval of others. His approach does not presuppose a convergence in thinking patterns and so is coincident with this study in this respect. Stevenson did not suggest any measurement. Moral loaded thinking patterns can always be evaluated by means of pq and PQ indices. Trials to influence others can be identified through tentative and reasoning qualities. However, measured indices do not reflect stages in morality. The approach in this study is open to amorality, like Stevenson's, since convergence of thinking patterns toward some justification or virtue is not presupposed.

Personality trait

In Chapter 6, a moderate but significant correlation between individual philosophical quality (pq) and *openness to experience* as personality trait is observed (cf. Cohen, 1977). This observation is relevant since it demonstrates that philosophical quality may be rooted in a stable underlying factor. This point will be developed in Section 9.6.

Sensitivity to pre-conceptual knowledge

Results of pq indices may also contribute to a deeper understanding of youngsters' sensitivity to pre-conceptual knowledge that cannot be explained directly by results of traditional cognitive measurements of philosophical thinking patterns. Twelve-year old Umi illustrates this (see Chapter 3). She expressed herself with respect to the philosophical quest for truth and beauty in Alexander Blok's poem: *Night*, *street*, *lamp*, *pharmacy*, displaying a sensitivity to the course of life events by her explicit question in that direction, while highlighting the meaning of the poem.

However, she produces a rather naïve explanation which leaves the observer doubting her understanding. Umi explains the perceived meaning of the poem by pointing to the need for light in the pharmacy when it is dark outside. Furthermore, she expresses a mystic feeling, serving as an analogy. It is doubtful whether Umi's stressed dark-light alternation in street and pharmacy can be explained as understanding a general chain of events like 'life-death' alternation. Also, no measurable link can be observed between this phenomenon and her mystical analogy reasoning of feeling like a flea. However, Umi's expressed thinking patterns obviously point to awareness of a cognitive notion but are unable to provide an adequate explanation. It seems that pq indicators can pick up ways of linking knowledge elements even before the child is able to put this knowledge in a rational line of reasoning. Umi's understanding can be sensed but not be quantified by traditional psychometric measurements, while assigning indicators, like indecisive thinking, tentative behaviour, openness and reasoning qualities to her utterances, record a philosophical quality. Indicator frequencies and indices for philosophical quality scored in this study do justice to thinking procedures, but not to capacities producing (philosophically) qualified contents.

6 *Philosophical quality as talent*

In Chapter 1 it was suggested that philosophical quality can be conceived as a *talent*: i.e., mental endowment with some genetic roots, involving rational beliefs, having the potential to be exploited in cooperation with the environment, and leading to valuable performance. High-level philosophical thinking is valued highly by society, and prodigies like Aristotle, Kant and Wittgenstein are often referred to as being talented or gifted. This is especially true when high-level performance is shown already at an early age.

The model of Ziegler and Heller (2000) and result of this study

In their meta-theoretical analyses of talent and giftedness, Ziegler and Heller (Ziegler & Heller, 2000; Heller, 2004) develop a model explaining exceptional achievement. In the Ziegler and Heller model, talent or giftedness is a core factor needed to explain successful performance, although alone it is not sufficient. Other conditions, specifically talent inducing personality traits and favourable environmental factors must be met in order to bring the individual to a critical state that can lead to excellent performance. This critical state is preceded by individual prior history and leads when facilitated by adequate developmental and learning processes, to increasing exceptional performance. The critical state may be attained at an early age, often before adolescence, dependent on the magnitude of environmental influences.

The results in this study can easily be fitted into the Ziegler and Heller model. In Chapter 2, an analysis of the main features of philosophy distinguished three pillars of philosophising. These were operationalised by five philosophical indicators. These indicators were found in thinking patterns of children of 11 years in age

and older. Indicators of philosophical quality were balanced and summarised as individual pq indices. Individual differences in these indices were found at all ages.

In older empirical studies of talent, giftedness was equated with general intellectual capacities. Modern intelligence and talent theories favour domain specific conceptions, exemplified by the multiple intelligence model developed by H. Gardner (1983, 1993, 1999). Gardner's theory of multiple intelligences identifies eight independent intelligences on the basis of distinct sets of processing operations applied in culturally meaningful activities: linguistic, logico-mathematical, musical, spatial, bodily/kinaesthetic, interpersonal, intrapersonal, and existential intelligence. Gardner's existential intelligence (in myth, art, science and philosophy) refers to the a separate intellectual domain focussing on thinking about the meaning of life and human existence. Again, the attempt to explain philosophical quality in this study can be compared with attempts to explain talent according to domains of intelligence, and seems to be based on hypothetical-deductive thinking. A divergent relationship between pq indices and intelligence was proven in this study. Likewise, philosophical quality could be linked to Hirst's exclusive fundamental cognitive category of 'philosophy' in domains of education (Hirst, 1980). Moreover, according to Ziegler and Heller, in addition to gifts one must take learning opportunity, a high level of motivation, and a supportive environment into consideration.

The Ziegler and Heller model begs attention to personal characteristics that are conducive to talent possibly on a genetic basis. In this study, it was demonstrated that philosophical quality is related to the personality trait 'openness to experience'. The role of environmental conditions for a talent to be functional is also emphasised by Ziegler and Heller. In the age range investigated, a linear relation between philosophical quality (pq) and age was not found. However, a clear link between pq indices and educational environment was detected. Moreover, the follow-up study reveals inhibiting influences that stem from the transition to schools of lower educational level. It appears that philosophical quality is open to exploitation in cooperation with the environment, and sensitive to developmental support or suppression from the environment (e.g., school and peers).

In review, it can be concluded that many features present in the Ziegler and Heller meta-theoretical talent model are replicated in the present research establishing philosophical quality. Therefore, the use of the term *talent* for this type of quality is considered justified.

7 Applicability of philosophical quality

Philosophical quality is quantified by means of pq and PQ indices. These indices allow comparison between performances of philosophically qualified thinking patterns. Moreover, they prove to be stable at individual levels and applicable to people in different degrees. To this end, such indices can be used as measures to discriminate people with respect to their philosophically qualified performance. This discriminative power has been assessed in philosophical discussions with 3 to 5 participants. This process can be extended for evaluation or selection purposes: philosophical essays

must be evaluated and mutually be compared; relevant selection criteria are required to select a qualified person for a professional job. Some professions that explore questionable or confusing issues are characterised by their need for philosophical quality, for example researcher scientists, journalists and inventors.

Evaluation criteria in philosophical education

The results of this study are derived from the philosophical discussions of four (minimum three and maximum five) participants. Discussions by just two persons often end up in philosophical debates. With more than five persons involved, the chance that all participate in an equal manner rapidly decreases. Moreover, rules for interference may begin to negatively impact direct expressions of thoughts. A simple application of tetralogues in the education system will not serve as a measurement to evaluate exam performance. To test the exclusive philosophical quality of papers, essays or other performances required to pass exams in philosophy, evaluation criteria have to be developed. The indices presented here for measuring philosophical quality may serve as a basis for such criteria if their outlines are adapted. Traditionally, evaluation of philosophical papers focused on convergent thinking patterns, qualified knowledge, language achievement, and supposed abilities to understand underlying principles. These qualities are valuable in the evaluation process, but they cannot be measured by the pq or PQ indices presented here. So long as exams test convergent thinking, pq and PQ indices will not be applicable.

However, the application of pq and PQ indices can contribute to a more complete evaluation of the philosophical contents because they reveal specific qualities identified by combinations of indicators on indecisive thinking, openness, tentative behaviour, epistemic position, and reasoning. Measurements might be adapted in two ways: 1) with respect to the unit of measurement; and 2) with respect to admitting discussion qualities to individual performances. In this study, the unit of measurement is equal to the unit of utterance, and is determined by the contribution of participant following. Since there are usually no other participants involved in writing papers and essays, an adequate unit of measurement may be established by introducing artificial breaks, for example by indentation or limited paragraphs. Secondly, papers and essays are individual performances that should identify interaction qualities between contrasting ideas and anecdotal qualities. Indices for individual philosophical quality could be extended and incorporated in a formula comparable with that of the index for group discussion. Such adaptations, including a pilot to test implementation and effects, could be realised in limited time-frame provided they are carried out by an expert with the full cooperation of the educational institutes and universities.

Selection of personnel

Some professions require exclusive philosophical qualities. These include scientific, journalistic and detective research that demand the ability to question systematically

and transcend smoothly through the domain of experience. Even inventing new devices or executive mechanisms requires these qualities. According to the British inventor and designer James Dyson (2001), the profession of 'inventor' is characterised by open-mindedness, sensitivity to the multifaceted nature of 'things', bringing questions up for discussion, persistency in research, and challenging authorities. Many of these requirements can be met through the philosophical quality assessed here. Indices for philosophical quality may provide selection criteria if applied to adequate utterances by potential candidates for these professions. Opportunities to urge candidates to express themselves may be generated through a discussion on philosophical topics with three other participants, or through a test exclusively designed for this purpose. Developing and testing a special measurement tool for applications might require several years of additional research. In contrast, implementing tetralogue discussions would not be very time consuming.

Developing a critical attitude: preventing from accepting things that 'go without saying'

Philosophical quality may also contribute to purposes of a more idealistic nature. One may appreciate philosophical quality in thinking patterns because of its appeal to an advanced thinking and of its perception of dealing with complex reality and experiences. It may also refer to a vague ideal of openness to the phenomena of a multi-layered reality. Philosophy may arise where different cultures meet each other. Qualities of sensitivity to ambiguity, vagueness and uncertainty, openness to the unknown and tentative behaviour may generate opportunities for people from different cultural backgrounds to meet in an unbiased way. Open discussions in politics would really be a blessing for society. Tentative interpretations of life events may evoke intellectual curiosity and prevent people from accepting things that 'go without saying'. Thinking in terms of what else is possible, and being open to unusual, unfashionable, uncommon or even 'sick' views may have social significance, serving emancipating purposes, and increasing the democratic content of society. This was the goal of earlier research into the external effects of the *Philosophy for Children* programme (Cwi 1976; Pálsson 1995; Schleifer and Poirier, 1995; Cam, 2000). While standard education and decorum presuppose some self-evident structures that undoubtedly serve an instrumental good, they also bear the risk of paralysing fundamental discussions about the potentially endless number of interpretations of reality. Above all, practicing philosophy requires the abandonment any pretension of knowing and attempting to find solutions for the confusing questions that arise during childhood already. In this respect, the PQ indices of group discussions may be a more adequate measure of philosophical quality than the evaluation of discussion topics, participation of different individuals, and debating winners and losers.

Perspectives to applied and fundamental research

Since all scientific research is driven by curiosity, it may be considered strange that many philosophical explorations by children and adolescents are disqualified

because they analyse and reason between imagination and measurable reality. The potential for such quests does not distinguish between fundamental and applied scientific research. Both fields of research are equally driven by curiosity and there is no reason to evaluate them unequally in society. Being sensitive to more than one dimension of obviousness, open to the unknown, and eager to shift from one to another domain of experience, will increase the philosophical quality of thinking patterns to obtain the afore mentioned ends. This study shows how to measure this philosophical quality as a talent.

References

American Educational Research Association (corporate author) (1999). Standards for Educational and Psychological Testing. Washington, DC: American Psychological Association National Council on Measurement in Education.

Ammers, E. van, Dekovic, M., Gerrits, L.A.W., Groenedaal, J.H.A., Hermanns, J.M.A., Meeuws, W.H.J., et al. (1998). Opvoeden in Nederland, schalenboek. Utrecht: Universiteit Utrecht.

Ardelt, M. (1998). Antecedents and effects of wisdom in old age: A longitudinal perspective on aging well. Research on aging, 22, 360-394.

Ardelt, M. (2004). Wisdom as expert knowledge system: A critical Review of a contemporary operationalization of an ancient concept. Human Development 47, 257-285.

Aristoteles (2000). Over dieren. J.M. van Ophuysen (Ed.) (R. Ferwerda, Trans.) Groningen: Historische uitgeverij.

Ashton, M.C, Goldberg, L.R., & Lee, K. (2004). A Hierarchical Analysis of 1,710 English Personality-Descriptive Adjectives. Journal of Personality and Social Psychology 2004, 87 (5), 707-721.

Bakeman, Roger, Gottman, & John M. (1997). Observing interaction: an introduction to sequential analysis. Cambridge UK: Cambridge University Press.

Baltes, P.B., & J. Smith (1990a). Weisheit und Weisheitsentwicklung: Prolegomena zu einer psychologische Weisheitstheorie. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 22 (2), 95-135.

Baltes, P.B., & J. Smith (1990b). Toward a theory of wisdom and its ontogenesis. In R.J. Sternberg (Ed.), Wisdom, its nature, origins, and development (87-120). New York: Cambridge University Press.

Baltes, P.B. & Staudinger U.M. (2000). Wisdom: A metaheuristic (pragmatic) to orchestrate mind and vitue toward excellence. American Psychologist, 55, 122-136.

Baltes, P.B. & Kunzmann, U. (2004). The Two Faces of Wisdom: Wisdom as a General Theory of Knowledge and Judgement about Excelence in Mind and Virtue vs. Wisdom as Everyday Realization in People and Products. Human Development (47), 290-299.

Bartsch, Renate. (1987). Norms of language: theoretical and practical aspects. London: Longman.

Berk, Laura E. (1997). Child Development. Massachusetts: Allyn and Bacon.

Birren, J.E. & Fischer, L.M. (1990). The elements of wisdom: overview and integration. In R.J. Sternberg (Ed.), Wisdom, its nature, origins, and development (317-333). New York: Cambridge University Press.

Blok, Alexander (1999). Nacht, straten, apotheek, lantaarn. In K. van het Reve: De meisjes van Zanzibar. Leiden: Plantage.

Boerwinkel, D.J. (2003). Het vormfunctieperspectief als leerdoel van natuuronderwijs. Utrecht: CD-β Press.

Bor, J. & Leeuw, K. L. van der (2003). 25 Eeuwen Oosterse Wijsbegeerte. Amsterdam: Boom.

Borsboom, D., Mellenbergh, G.J., & Heerden, J. van (2004). The Concept of Validity. Psychological review 111 (4), 1061-1071.

Brugman, G.M. & Dudink, A.C.M. (1976). De vraag naar creativiteit. Meppel: Boom.

Brugman, G.M. (2000). Wisdom: Source of narrative coherence & eudaimonia. Delft: Eburon.

Brugman, G.M. & Dudink, A.C.M. (2002). Intelligentie, creativiteit en wijsheid. In J.J.F. Schroots (Ed.), Handboek Psychologie van de Volwassen Ontwikkeling & Veroudering. Assen: Koninklijke Van Gorcum.

Brugman, G.M. (in press). Wisdom and Aging. In J.E. Birren & K.W. Schaie (Eds.), Handbook of Psychology of Aging (6th ed.). San Diego: Academic Press.

Cam, P. (2000). Philosophy and Freedom. Thinking, 15 (1), 10-13.

Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

Cebas, E. & Moríyón, F.G. (2003). What we know about research in Philosophy for Children. (unpublished)

Chandler, M.J. and Holliday, S.G. (1990). Wisdom in a postapocalyptic age. In R.J. Sternberg (Ed.), Wisdom, its nature, origins, and development (121-142). New York: Cambridge University Press.

Chapman, Michael (1999). Constructive Evolution, Origins and development of Piaget's thought. New York: Cambridge University Press.

Chomsky, Noam (1957). Syntactic Structures. Den Haag: Mouton.

Clark, E.V. (1979). The ontogenesis of meaning. Wiesbaden: Athenaion.

Clark, E.V. (1983). Meaning and concept. In P.H. Mussen (Ed.), Handbook of childpsychology: Cognitive development. New York: Wiley.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20 (1), 37-46.

Cohen, J. (1968). Weighed Kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychological Bulletin 70 (4), 213-221.

Cohen, J. (1977). Statistical power analysis for the behavioural sciences. New York: Academic Press.

Crohnbach, L.J., & Mehl, P.E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281-302.

Cwi, D. (1976). Competency, the Underprivileged, and Elementary School Philosophy. Metaphilosophy 7 (1), 76-79.

D'Andrade, R. (1989). A folk model of mind. In D. Holland and N. Quinn (Eds.), Cultural Models in Language and Thought. New York: Cambridge University Press.

Dawes, R. M. (2001). Everyday Irrationality. Oxford: Westview Press.

Deleuze G., & Guattari, F. (1994). What is Philosophy? London: Verso.

Denett, D. (1993). Het bewustzijn verklaard. In: W. Kayzer (Ed.), Een schitterend ongeluk. Amsterdam: Contact.

Derrida, J. (1967). L' ecriture et la difference. Paris: Seuil.

Descartes, R. (1979). Over de methode. Meppel: Boom.

Dewey, J. (1956). The Child and the Curriculum and The School and Society. Chicago: University of Chicago Press.

Dewey, J. (1966). Democracy and Education. An Introduction to the Philosophy of Education. New York: The Free Press.

Dixon, Roger A. & Lerner, Richard M. (1992). A history of Systems in Developmental Psychology. In Marc H. Bornstein & Michael E. Lamb, Developmental Psychology: An Advanced Textbook. Hillsdale, New Jersy: National Institute of Child Health and Human Development.

Drenth, P.J.D., & Sijtsma, K. (1990). Inleiding in de theorie van de psychologische test en zijn toepassingen. Houten: Bohn, Stafleu en Van Loghum.

Dyson, J. (2001). A history of great inventions. New York: Carroll & Graf Publishers.

Elbers, E.P.J.M. (1993). Leren door interactie. Groningen: Noordhoff.

Erikson, E.H. (1968). Identity, Youth and Crisis. New York: W.W. Norton.

Evers, A. (2001). Beoordelingsschema voor tests en studietoetsen. Unpublished report: Commissie testaangelegenheden van het NIP (COTAN).

Frank, A. (1997). Het Achterhuis, dagboek van Anne Frank. Amsterdam: Bert Bakker.

Freese, H.L. (1990). Kinder sind philosophen. Berlin: Quadriga.

Gardner, H. (1983). Frames of Mind: The theory of multiple intelligences. New York: Basic Books.

Gardner, H. (1993). Creating minds. New York: Basic Books.

Gardner, H. (1999). Intelligence reframed: Multiple intelligences for the 21th century. New York: Basic Books.

Gibson, J.J (1979). The ecological approach to visual perception. Boston: Houghton Mifflin.

Glaser, B.G., & Strauss, A.L. (1967). The discovery of grounded theory: Strategies for qualitative research. Chicago: Aldine.

Guilford, J.P. (1988). Some Changes in the Structure-of-Intellect Model. Educational and Psychological Measurement, 48, 1-4.

Hadot, Pierre (2003). Filosofie als een manier van leven. Amsterdam: Ambo.

Hartmann, D. P. (1977). Considerations in the choice of interobserver reliability estimates. Journal of Applied Behavior Analysis, 10, 103-116.

Heller, K.A. (2004). Identification of gifted and talented students. Psychological Science, 46 (3), 302-323.

Hendriks, A.A.J. (1997). The construction of the Five-Factor Personality Inventory FFPI. Groningen: Rijks Universiteit Groningen.

Herodotus (1995). Historiën (historiae). In H. van Dolen, Het verslag van mijn onderzoek. Nijmegen: SUN.

Hirst, P.H. (1980). Liberal education and the nature of knowledge. In R.S. Peters, Philosophy of Education (pp. 87-111). London: Oxford University Press.

Hirst, P.H. (1983). Educational theory and its foundation disciplines. London: Routledge.

Hoekstra, H.A., Ormel, J., & Fruyt, F. de (1996). NEO PI-R NEO FFI, Big Five Persoonlijkheidsvragenlijst. Lisse: Swets & Zeitlinger B.V.

Hofstadter, D. (1985). Gödel, Escher, Bach. Amsterdam: Contact.

Holliday, S.G., & Chandler, M. J. (1986). Wisdom: explorations in adult competence. Basel: Karger.

Husserl, E. (1968). Logische Untersuchungen, Untersuchungen zur Phänomenologie und Theorie der Erkenntnis. Tübingen: Max Niemeyer Verlag.

Husserl, E. (1996). Die Krisis der europäische Wissenschaften und die transzendentale Phänomenologie. Hamburg: Felix Meiner Verlag.

Jackson, S. (1995). Adolescent Identity Development and Social Context: Towards an Integrative perspective. In A. Oosterweegel & R.A. Wicklund (Eds.), The Self in European and North American Culture: Development and Proceses. Dordrecht: Kluwer Academic Publishers.

Kant, I. (2004). Kritiek van de zuivere rede. Amsterdam: Boom.

Kayzer, W. (1993). Een schitterend ongeluk: Inleiding. Amsterdam: Contact.

Kekes, J. (1983). Wisdom. American Philosophical Quarterly 20 (3), 277-287.

Kitchener, K.S. (1983). Cognition, metacognition, and epistemic cognition. Human Development, 26, 222-232.

Kitchener, K.S. (1990). Wisdom and Reflective Judgment: Knowing in the face of uncertainty. In R.J. Sternberg (Ed.), Wisdom, its nature, origins, and development. Cambridge University Press.

Kitchener, K.S., Lynch, C.L., Fischer, W. & Wood, P.K. (1993). Developmental range of reflective judgment: The effect of contextual support and practice on developmental stage. Developmental Psychology, 29, 893-906.

Kohlberg, L. (1971). From is to ought. In Th. Mischell, Cognitive development and epistemology. New York: Academic Press.

Kohlberg, L. (1981). Essays on moral development. San Francisco: Harper and Row.

Koops, W. (2001). Agressieve ontwikkelingspsychologie. Universiteit van Utrecht.

Kuhn, Thomas (1975). The structure of scientific revolutions. Chicago: University of Chicago Press.

Laak, J.F. ter, & Goede, M.P.M. de (2003). Psychologisch Diagnostiek. Lisse: Swetz & Zeitlinger.

Leeuw, K.L. van der (1994). Het Chinese Denken. Amsterdam: Boom.

Leeuw, K.L. van der (2005). Philosophical Dialogue and the Search for Truth. Thinking, the Journal of Philosophy for Children 17 (3), 17-23.

Lieshout, C.F.M. van, Heymans, P.G. (Eds.) (2000). Developing Talent Across the Life Span. Hove, East Sussex: Psychology Press.

Linley, P.A. (2003). Positive Adaptation to Trauma: Wisdom as Both Process and Outcome. Journal of Traumatic Stress, 16 (6), 601-610.

Lipman, M. (1976). Philosophy for Children. Metaphilosophy 7 (1), 17-39.

Lipman, M. (1985). Thinking Skills fostered by Philosophy for Children. In S. Chipman, J. Segal and R. Glaser (Eds.), Thinking and Learning Skills: relating instruction to basic research, Vol. 1. Hillsdale, NJ: Erlbaum.

Lipman, M. (1991). Thinking in Education. Cambridge U.K.: Cambridge University Press.

Lipman, M., Sharp, A.M. & Oscanyan, F. (1977). Philosophy in the Classroom. West Caldwell, NJ: IAPC.

McCrea, R.R., & Costa, P.T. (1990). Personality in Adulthood. New York: Guilford Press.

Matthews, Gareth B. (1980). Philosophy and the young child. Cambridge, MA: Harvard University Press.

Matthews, Gareth B. (1994). The Philosophy of Childhood. Cambridge, MA: Harvard University Press.

Matthews, Garreth B. (1995). Filosofie van de kinderjaren. Rotterdam: Lemniscaat.

Meacham, J.A. (1990). The loss of wisdom. In R.J. Sternberg (Ed.), Wisdom, its nature, origins, and development (pp 181-211). New York: Cambridge University Press.

Merleau-Ponty (1967). La Phenomenologie de la perception. Paris: Gallimard.

Messick, S. (1989). Validity. In R.L. Linn (Ed.), Educational measurement (13-103). Washington, DC: American Council on Education and National Council on Measurement in Education.

Meyer, G.J., Finn, S.E., Eyde, L.D., Kay, G.G., Moreland, K.L., Dies, R.R., et al. (2001). Psychological Testing and Psychological Assessment. A Review of Evidence and Issues. American Psychologist, 56 (2), 128-165.

Morris, G.H., White, C.H. & Iltis, R. (1994). 'Well, Ordinarily I Would, But', Reexamining the Nature of Accounts for Problematic Events. Research on Language and Social Interaction, 27(2), 123-144. Lawrence Erlbaum Associates, Inc.

Nagel, T. (1974). What is it like to be a bat? The Philosophical Review 83 (4), 435-450.

Nagel, T. (1990). Wat betekent het allemaal? Amsterdam: Bert Bakker.

Nelson, L. (1970). System der Philosophische Ethik und Pädagogik. Hamburg: Felix Meiner Verlag.

Nelson, L. (1994). In J. Kessels (Ed.) De Socratische methode. Amsterdam: Boom.

Nietzsche, F. (1999). De vrolijke wetenschap. Amsterdam: De Arbeiderspers.

Niklasson, J., Ohlsson, R. & Ringborg, M. (1995). Evaluating Philosophy for Children. Thinking, The Journal of Philosophy for Children 12 (4), 17-23.

Nozick, Robert (1981). Philosophical Explanations. Cambridge, MA: Harvard University Press.

Nunnaly, J.C. & Bernstein, I.H. (1994). Psychometric Theory. New York: McGraw Hill Inc.

Olthof, T. & Brugman, D. (1994). Het ontstaan van moreel besef. Lisse: Swets & Zeitlinger.

Pálsson, H. (1995). We think more than before about others and their opinions. Thinking, The Journal of Philosophy for Children 12(4), 24-29.

Peirce, C.S. (1992). Reasoning and the logic of things. (K. Ketner, ed.). Cambridge, MA: Harvard University Press.

Pels, T. (Ed.). (2000). Opvoeding en Integratie, een vergelijkende studie van recente onderzoeken naar gezinsopvoeding en pedagogische afstemming tussen gezin en school. Assen: Van Gorcum.

Pennebaker, J.W., Mehl, M.R., & Niederhoffer, K.G. (2003). Psychological Aspects of Natural Language Use: Our words, Our Selves. Annual Revews Psychology, 54, 547-577.

Peters, R.S. (1980): The Justification of Education. In R.S. Peters (Ed.), The philosophy of education (pp. 239-268). Oxford.

Philipse, H. (2004): Wetenschap en wereldbeeld. Universiteit van Utrecht.

Piaget, J. (1973). The Child's conception of the world. London: Granida.

Piaget, J. & Inhelder, B. (1974). The Child's Construction of Quantities: Conservation and Atomism. London: Routledge & Kegan Paul.

Raad, B. de (1994). An expedition in search of a fifth universal factor: Key issues in the lexical approach. European Journal of Personality, 8, 229-250.

Raven, J.C., Court, J.H., & Raven, J. (1977). Standard Progressive Matrices. London: H.K. Lewis & Co. Ltd.

Riegel, K.F. (1973). Dialectic Operations: The final Period of Cognitive Development. Human Development 16, 346-370.

Rondhuis, N.Thecla W. (1990): De zin van onzin. Doctoraal scriptie, Universiteit van Amsterdam (unpublished).

Rondhuis, N.Thecla W. (1994): Ik dacht bij mezelf… [TV programme]. Hilversum: IKON

Rondhuis, N.Thecla W. (1994). Filosoferen met kinderen. Rotterdam: Lemniscaat.

Rondhuis, N.Thecla W. & Leeuw, K.L. van der (2000): Performance and Progress in Philosophy: An Attempt at Operationalisation of Criteria. Teaching Philosophy 23 (1), 23-42.

Rondhuis, N.Thecla W. (2001). Jong en Wijs. Amsterdam: Piramide.

Rondhuis, N.Thecla W. (2001, October 6). Kind & Poëzie. Trouw, pp.6-7. Amsterdam.

Rosch, E. (1978). Principles of categorisation. In E. Rosch (Ed.), Cognition and categorisation. Hillsdale, NJ: Erlbaum.

Russell, B. (1974). History of Western Philosophy. London: George Allen and Unwin.

Saint-Exupery, Antoine de (1980). De Kleine Prins. Rotterdam: Ad Donker.

Santi, M. (1993). Philosophizing and Learning to Think: Some Proposals for a Qualitative Evaluation. Thinking, The Journal of Philosophy for Children 10(3), 16-22.

Santi, M. (1995). Ragionare con il discorso; Il pensiero argumentativo nelle discussioni in classe. Firenze: La Nuova Italia.

Saucier and Goldberg (1996). Evidence for the Big Five in analyses of familiar English personality adjectives. European Journal of Personality, 10, 61-77.

Saussure, Ferdinand de (1976). Cours de linguistique générale. Paris: Payot.

Schleifer, M., & Poirier, G. (1995). The Effect of Philosophical Discussions in the Classroom on Respect for Others and non-Stereotypic Attitudes. Thinking, The Journal of Philosophy for Children 12 (4), 32-34.

Schleifer, M., & Courtemanche, L. (1995). The Effect of Philosophy for Children on Language Ability. Thinking, The Journal of Philosophy for Children 12 (4), 30-31.

Searl, J.R. (1983). Intentionality. Cambridge: Cambridge University Press.

Sellars, W. (1963). Science, Perception and Reality. Londen: Routledge & Kegan Paul.

Selter, Christoph & Spiegel, Harmut (1997). Wie Kinder rechnen. Leipzig: Ernst Klett Grundschulverlag.

Shipman, V. (1983). The New Jersey Test of Reasoning Skills. New Jersey: Totowa Board of Education.

Simonton, D.K. (1988). Scientific genius: A psychology of science. Cambridge: Cambridge University Press.

Simonton, D.K. (1990). Creativity and wisdom in Aging. In J.E. Birren & K.W. Schaie (Eds.), The Handbook of the psychology of Aging (320-329). San Diego: Academic Press.

Simonton, D.K. (1994). Greatness. Who makes history and why. New York: The Guilford Press.

Simonton, D.K. (2003). Scientific Creativity as Constrained Stochastic Behavior: The Integration of Product, Person, and Process Perspectives. Psychological Bulletin, 129 (4), 475-494.

Skinner, B.F. (1957). Verbal behavior, New York: Appleton-Cetury-Crofts.

Slade, Chr. (1992). Creative and Critical Thinking. An Evaluation of Philosophy for Children. Analytic Teaching, 13(1), 25-36.

Staudinger, U.M., Lopez, D. & Baltes, P.B. (1997). The psychometric location of wisdom-related performance: Intelligence, personality and more? Personality and Social Psychology Bulletin, 23, 1200-1214.

Sternberg, R. J. (1990). Wisdom and its relations to intelligence and creativity. In R.J. Sternberg (Ed.), Wisdom, its nature, origins and development. Cambridge University Press.

Sternberg, R.J. & Lubart, T.I. (1996). Investing in creativity. American Psychologist, 51, 677-688.

Sternberg, R. J. (2000a). Wisdom as a form of giftedness. Gifted and Child Quarterly, 44, 252-260.

Sternberg, R. J. (2000b). Successful intelligence: A unified view of giftedness. In C.F.M. van Lieshout & P.G. Heymans, P.G. (Eds.). Developing Talent Across the Life Span. Hove, East Sussex: Psychology Press.

Sternberg, R. J. (2003). Wisdom, intelligence, and creativity synthesized. New York: Cambridge Universty Press.

Sternberg, R.L. (2004). Words to the Wise about Wisdom? Human Development 47, 286-289.

Stevenson, Ch. L. (1963). Facts and Values. Studies in ethical analysis. New Haven, CT: Yale University Press.

Stewart, Susan (1979). Nonsense: aspects of intercontextuality in folklore and literature. Baltimore: John Hopkins University Press.

Stokhof, M. (2001): Het goede leven en het virtuele zelf. In T. Koetsier, K. Vintges & H. Schwab (Eds.), Word ik van filosofie een beter mens? Budel: Damon.

Sugarman, L. (2002). Frameworks, Accounts and Life-Span Development. London: Psychology Press.

Turner, S.M. & Matthews, G. B. (Eds.) (1998). The Philosopher's Child. University of Rochester Press.

Vandaele, V. (1996). Filosoferen met kinderen: Een evaluatie van enkele lopende projecten in Gentse Freinet-scholen. Doctoraal scriptie, Universiteit van Gent (unpublished).

Verhoeven, C.W.M. (1973). Inleiding tot de verwondering. Bilthoven: Ambo.

Verhofstadt-Denève, L., Van Geert, P., & Vyt, A. (1995). Handboek Ontwikkelingspsychologie. Houten: Bohn, Stafleu en Van Loghum.

Vygotsky, L.S. (1964). Denken und Sprechen. Berlin: Akademie Verlag.

White, D.A. (2003). Philosophy and Theory in the Study of Gifted Children. Roeper Review, 26 (1), 16-19.

Wittgenstein, Ludwig (1969). Tractatus logico philosophicus. Frankfurt am Main: Suhrkamp.

Wittgenstein, Ludwig (1976). Filosofische onderzoekingen. Meppel: Boom.

Yule, Sandy, & Glaser, Jen (1994). Classroom Dialoque and the Teaching of Thinking. Melbourne: Institute of Education Research Report.

Ziegler, A. & Heller, K. A. (2000). Conceptions of Giftedness from a Meta-Theoretical Perspective. In K.A. Heller, F.J. Mönks, R.J. Sternberg, & R.F. Subotnik (Eds.), International Handbook of Giftedness and Talent. Amsterdam: Elsevier.

Appendix I

Samples of philosophical discussions (Dutch Version)

Begin van de Donau

Steve: Ik vind dat we niet kunnen weten wat de oorsprong van de Donau is. We kunnen alleen kijken naar monsters van het water of van de grond: welke de oudste is.

Inez: Misschien is één van die zijrivieren wel de eigenlijke Donau.

Prisca: Waarom heet de Breche niet Mindel of de Mindel niet Breche en die niet Donau?

Rein: Mensen hebben dat zo genoemd. Maar misschien moet je daarover nadenken.

Steve: De Donau zou best wel eens ouder kunnen zijn dan de mensen.

Prisca: Hoe weet je dat?

Steve: Dat lees je in boeken.

Inez: Maar nou weten we nog steeds niet waar het begin van de Donau is. Als je daarover gaat nadenken, dan hoeft wat je denkt toch niet altijd waar te zijn? Je kan wel proberen gelijk te hebben, je kan echt denken: ik heb gelijk.

Steve: Wat waar is, is toch gewoon een gokje. En al is het gokje waar, dan was het maar een gokje.

Inez: Oké. Die ene zijtak is de langste. Die kan dus de meeste regen vangen en ik denk dat die dan ook de oudste is. Nou, dat kan wel. Want de oudste vormt de langste weg omdat die daarvoor de meeste tijd heeft gehad. Kunnen we niet kijken welke tak het beste met de Donau meevormt? Over dat oudste water: als dat er is, ligt het al lang in de Zwarte Zee en weet je nooit van welke rivier het komt.

Jacob: Het kan toch ook regen zijn geweest?

Inez: Waterdruppeltjes in de Zwarte Zee kunnen opnieuw verdampt zijn, en nu in de Atlantische Oceaan …

Jacob: Al dat water moet uit regen zijn gemaakt. En als water nou water is … Water heeft toch gewoon geregend? Dan heeft het toch heel lang geregend? Dan is het toch gewoon allemaal hetzelfde water? Volgens mij is gewoon alles Donau. Maar waar komt water dan vandaan?

Niets. Is niets misschien toch iets?

Jochem: 'Niets' is onzichtbaar. Als dat in een fles zit, dan zie je dat niet.

Casper: Je ziet dat daar niets in zit. Dus je ziet wat.

Rémy: Dan is 'niets' dus aan de ene kant niet iets en aan de andere kant weer wel. Want we praten wel over iets.

Timo: Nee! Niets is niet iets.

Berend: Misschien bedriegen je ogen je wel. Misschien is er wel iets, maar zie je het niet.

Rémy: Maar als je het niet kan zien, hoe weet je dan dat het er is?

Casper: Ja! Misschien staat dáár wel een kerk.

Berend: Maar als je ogen je bedriegen, kun je altijd nog tegen die kerk oplopen.

Timo: Daar kan geen kerk staan. Een kerk is door mensen gemaakt en die moeten dat zien.

Wouter: Je kan een soort sixth sense effect hebben, zo van: 'I see that people', het is niks, maar je ziet ze wel.

Berend: Hoe weet je nou of de dingen die je wel ziet, maar waarvan je zegt dat ze er niet zijn, er echt niet zijn?

Wouter: Nou, omdat ze er voor de rest van de mensen niet zijn.

Jochem: Niets kan ook zijn wat je niet kent. Dan zeggen ze: 'astromolieteblub', en dan denk jij: 'dat ken ik: dat is helemaal niets'. En verder kan 'niets' nog iets betekenen, bijvoorbeeld als je op de rommelmarkt 'niets' vindt.

Wouter: Kijk, als je 'niets' proeft, dan proef je wel iets, namelijk 'niets'. 'Niets' op de rommelmarkt krijgt misschien straks nog een naam. Dan gaat iedereen opeens 'niets' verkopen.

Casper: Maar je zegt toch niet: lekker zeg, dat 'niets'.

Wouter: 'Niets' is wat wij niet kunnen benoemen.

Marit: Maar bij 'niets' merk je dus niets. Nergens vind je 'niets'. Dus de vraag klopt niet. Wat is nou het verschil tussen niets en iets? Niets? Er is altijd iets. Wie kan mij nou vertellen wat 'niets' en wat 'iets' is?

Timo: Als iets geen naam heeft, dan vind ik niet dat het niets is. Op een gegeven moment hadden we de sterren nog niet ontdekt, toen hadden we nog nooit sterren gezien. Bestaan er dan geen sterren?

Jochem: Maar als je iets niet hebt gezien en het heeft wel een naam, kun je niet zeggen dat het iets is. Je bent er niet zeker van.

Berend: Als onderzoekers een nieuwe diersoort vinden, nemen ze er foto's van en laten ze die aan iedereen zien. Maar dan heeft hij nog geen naam.

Wouter: Als een baby nog nooit zuurkool heeft gezien of geroken, dan is het toch nog steeds zuurkool?

Timo: Maar dat weet je dan niet.

Marit: Maar wat zit er dan in een fles met 'niets'? Onzichtbare zuurkool? Het moet iets zijn.

Jochem: Misschien is 'niets' een groep van voorwerpen die we niet zo goed kennen en die 'niets' heet. Op een rommelmarkt zie je vazen, een trompet, oude dingen … Maar die zoek je niet. Het woord 'niets' kennen we wel en het betekent wel iets, maar we kunnen het niet zien. We hebben het ook nog nooit gezien. Dit is echt een heel lastig onderwerp en ook heel raar.

Kunnen de landen op de aardbol veranderen?

Rex: Landen kunnen wel en toch ook niet veranderen. Landen kunnen veranderen op kaarten, maar niet in het echt. De grond kan veranderen. Na een vulkaanuitbarsting kan lava weer nieuw land worden. En ... de techniek ontwikkelt zich ook weer verder. Zo is het gekomen dat wij nu denken dat de wereld rond is in plaats van plat. Dus, misschien zijn er later ook weer nieuwe inzichten.

Najia: Je kunt Marokko ook groter maken. Als iedereen Marokko groter dan Nederland wil hebben, misschien omdat er in Duitsland, in Nederland, in ... weet ik het, overal veel Marokkanen wonen, die weer naar Marokko willen gaan.

Lisa: Hoe kun je nou Marokko groter maken? Meer huizen ... ja, maar dan wordt Marokko toch niet groter?

Willem: Nou, met zand ... maar als je zand ergens anders vandaan haalt, dan is dat daar weer weg.

Esther: Ja, waar haal je dat zand dan vandaan? Uit de Sahara-woestijn? Maar dan wordt die weer kleiner.

Ramona: Dan wordt de Sahara niet kleiner. Het gaat alleen om die zandberg. Je raakt zand kwijt, maar dan wordt de Sahara toch niet kleiner.

Rex: Later, heb je misschien een machine, die de zwaartekracht kan opheffen. En dan kun je als het ware een land weghalen en op Marokko zetten. Maar dan heb je weer een probleem: dan is dat stukje land geen Marokko meer.

Tjalling: Maar als Marokko nou bijvoorbeeld oorlog gaat voeren met Algerije en als die wint, dan heeft die ook weer land erbij.

Ramona: Nou, Najia zei van dat mensen het moeten willen, maar ik denk niet dat die Algerijnen dat willen.

Lisa: Je kan met oorlog voeren de wereld wel verplaatsen. Als nou twee landen, bijvoorbeeld Spanje en Italië, oorlog met elkaar gaan voeren. En Italië wint een stuk van Spanje, Spanje een stuk van Italië en dat gaat een hele tijd door, zodat op een gegeven moment Italië Spanje gewonnen heeft en Spanje Italië. Dan ligt Italië in Spanje en ...

Rex: Maar kijk, Lisa, jij zegt dat ze dan verwisseld zijn, maar eigenlijk is het alleen maar de naam. Ik kan Nederland Mars noemen, maar dat wil nog niet zeggen dat Nederland Mars is.

Lisa: Maar Nederland is toch ook Nederland geworden en Italië was toch eerst denk ik ook wat anders? Een land kan toch nú ook anders worden?

Rex: Ja, dat geloven wij. Maar misschien is het wel heel anders.

Tjalling: Met oorlogen kun je het ene land groter maken, maar wordt het andere kleiner.

Rex: Eigenlijk kunnen we dingen wel groter maken. Want als je bijvoorbeeld een mier onder een microscoop houdt, dan wordt die ook groter voor je zien. Dan moet je alleen nog de formule weten hoe je dat verder, in het echt zeg maar, groter maakt.

Oneindigheid

Andrea: Je kunt in een trein altijd maar doorrijden langs stadjes en zo. Je blijft altijd maar rails zien. Dat stopt niet. Wanneer stopt oneindigheid eigenlijk?

Ron: Er komt altijd een einde, iedere trein komt bij een eindstation.

Paco: Maar die rails gaat wel gewoon door.

Jerry: Bij snelwegen heb je de A2, de A9 en ... Als je dan verdergaat wordt het bijvoorbeeld A3. Nou dan gaat het eindpunt van de A2 over op de A3.

Paco: Dan krijgt die weg alleen een andere naam.

Jerry: Ja, dus is het een andere weg. Want kijk, op de A2 mag je 120 km en op de A3 mag je maar 100. Dan is het toch een andere weg?

Umi: Als die A2, A3 en zo, dezelfde kant uitgaan, kun je tot in het oneindige doorrijden. Maar als je bij een bocht komt, is dat het einde van die rechtdoor gaande weg.

Jerry: Ja, kijk bijvoorbeeld de Huygenstraat is met klinkers. Die gaat dan over in het Tesselschadeplein. Dat is asfalt. Dan is het een ander wegdek en dus een andere straat.

Andrea: Oneindigheid is dat je maar door blijft gaan. Er komt geen einde aan. Dat is ook zo met rondjes rijden. Voor mij bestaat oneindigheid in ieder geval wel. Als een weg een andere naam krijgt is het niet meteen een andere weg. Snelwegen zijn gewoon één netwerk van allemaal stukjes asfalt die over de hele wereld leiden.

Paco: Misschien kun je wel naar het verdwijnpunt reizen, met genoeg stroom, zuurstof en misschien ben je wel onsterfelijk. Misschien kom je dan ergens uit waar je niets weet. Dus misschien is er wel oneindigheid, maar ook weer niet, want je weet dat dan niet.

Ferhat: Het is een illusie. Waarom noem je dat oneindig? Noem het oneindigend.

Poëzie

Umi: Toch ... het gedicht gaat een beetje over 'wat heeft het leven voor een zin?' Het leven draait gewoon door. Alles staat met elkaar in verband. Vroeg of laat komt de apotheek bij de lantaarn. In de apotheek heb je licht en de lantaarn geeft licht. Het gedicht zet je aan het denken: 'zou het waar zijn?', 'Kan iemand weer iets anders worden?' Er zijn gedichten die niet 'rijmen' maar het blijven gedichten, want ze geven je een bepaald 'indenken': een gedachte die jou ook kan overkomen. Je kan wel eens iets geks wensen. Ik heb wel eens gewild dat ik een vlo was. Hoe zou dat zijn? Het klinkt zo logisch, maar het zal allemaal heel anders zijn: alleen maar grote dingen om je heen, niet naar school, heel vreemd. Dat kan een gedicht zijn.

Verkeerd stemmen

Marlon: Ik vind het raar dat iemand die niet goed kan denken, wel kan stemmen. Misschien stemt hij wel verkeerd!

Desi: Iemand die niet goed kan nadenken is toch ook een mens? Het is toch geen alien?

Nina: Iedereen boven de 18 mag stemmen, of je nou wel of niet goed kan nadenken.

Jos: Stel hij moet stemmen voor 'de wereld moet meteen opgeruimd worden' of 'de wereld moet niet opgeruimd worden', en hij drukt op het verkeerde knopje ...

Nina: Wat is nou het verkeerde knopje?

Lucy: Dat kun je hem toch uitleggen

Marlon: Het kan toch zijn dat hij iets verkeerds kiest, iets dat hij niet wil.

Lucy: Het is nooit verkeerd. Misschien weet hij niet wat hij nou doet, maar verkeerd ...

Nina: Het knopje waarop hij drukt is zíjn keuze, zíjn gedachte.

Jos: Andere mensen, die wel goed kunnen denken, kunnen ook niet goede gedachtes hebben.

Lucy: Het blijft zijn mening.

Marlon: Natuurlijk niet. Hij weet niet eens wat hij stemt. Hij weet ook zijn mening niet.

Desi: Misschien kan zijn moeder zeggen: druk maar op dat knopje dat ik het goede vind.

Marlon: Geeft die stem van hem dan zíjn mening? Nee toch! Hij drukt gewoon op het knopje omdat hij dat een mooi knopje vindt. En als hij dan voor het verkeerde kiest?

Lucy: Er is geen verkeerde. Want er komt nooit een vraag van: zullen we het leven op aarde doodmaken.

Nina: Er is wel goed en verkeerd. Maar niet in stemmen. Er zijn alleen maar partijen die willen winnen en er worden nooit belangrijke vragen gesteld, zoals van of we het leven op aarde moeten laten doodgaan of zo.

Lucy: Maar kijk, als je zegt dat sommige mensen niet kunnen stemmen, kun je net zo goed zeggen dat sommige mensen niet naar de kermis mogen.

Jos: Ja, dat is een goede, bijvoorbeeld omdat iemand er niet goed bij nadenkt dat hij in de achtbaan kotsmisselijk kan worden.

Nina: Als hij er toch ingaat, is dat toch zijn eigen keuze? Daar ga je toch niet dood van?

Jos: Nee, van verkeerd stemmen ook niet.

Nina: Goed en verkeerd zijn anders in dit voorbeeld.

Desi: Nee, het is hetzelfde.

Marlon: Maar stemmen heeft natuurlijk wel iets met nadenken te maken en niet met welke knop je het mooist vindt.

Lucy: Professors kunnen ook verkeerd stemmen. Iedereen kan fouten maken.

Desi: Mijn oma vergeet ook altijd alles.

Nina: Eigenlijk is het onzin om te zeggen dat het alleen boven de 18 mag. Op je tiende kun je misschien ook wel goed nadenken.

Jos: Of je het kunt weet je alleen zelf, eigenlijk.

Marlon: Ja, welke knop vind je het mooist. Misschien kies je dan wel goed. Jij vindt misschien dat goed stemmen afhangt van journaal kijken. Nou dat joch van 'Goede Tijden' is 38, maar gedraagt zich als zes. Als iemand van 38 een keiharde klap op zijn kop heeft gekregen en zich gedraagt als iemand van zes, mag hij dan stemmen? Ik vind het raar. Ben je dan voor jezelf 18 of voor andere mensen? Gaat het nou om je gedachtes, of om je lijf?

Appendix II

Characteristics of indicators on three levels (Dutch)

Woorden **Uitdrukking** **Uitdrukkingswijzen**	**Betekenisaspecten** **Parafrasering** **Illustraties**	**Scores** **Waarden**
Lijken	Ambiguïteit	Idt
Schijnen	Aanvoelen van nattigheid	Idt
	Gekoppeld aan de eerste persoon	Ep
	Beschrijving van een gelijkenis	-
Als het ware	Ambiguïteit	Idt
	Proberend	Te
	Stopwoord	-
Alsof, net als	Associatie, metafoor	Te
Schijnbaar	Ambiguïteit	Idt
Blijkbaar	Relativerend, scepsis	Idt & Ep
Kennelijk	Tentatief	Te
	Functie in redenering	Re
Blijken	Functie in redenering	Re
	Vaststellen (neutraal beschrijvend)	-
Noemen, Heten	Ambiguïteit (uitgezonderd eerste persoon singular)	Idt
Beschouwen als,	Ik noem ... (eerste persoon singular)	Ep
Ze zeggen ...	Constatering, vaststelling	-
Het is maar net hoe je het bekijkt/interpreteert/ beschouwt	Relativerend	Idt
"Tussen aanhalingstekens"	Relativerend	Idt
Alsof	Vaag, ambigu	Idt
	Vergelijking in dienst van redenering	Re
Eigenlijk	Ambiguïteit, grensgeval	Idt
Feitelijk	Voorzichtigheiduitdrukking	-
In principe, in het echt	Proberend	Te
Op zich	Stopwoord	-
Echt	Werkelijk, waar	-
...echt ...	Versterking van betekenis	-
In het echt	Ter onderscheid van niet-echt	Idt
	Hakkelend, uitstel van oordeel	Te
Je hebt geen echt bewijs	Onzekerheid, relativiteit	Idt
(niet) iedereen ziet het	Onzekerheid, relativiteit	Idt
Een beetje, half, vaag	Ambiguïteit	Idt
Soort van	Onderscheidend	Idt
Manier van	Proberend	Te
Min of meer	Voorzichtigheidsuitdrukking	Te
Sommige ... ook	Stopwoord	-
Moeilijk	Niet weten, constateren van ambiguïteit	Idt
	Zich verwonderend	Op

Je moet het zo zien	Arbitrair karakter van oordeel	Idt
Het valt bijna niet te bewijzen	Onzekerheid	Idt
Ander(e)	Constateren van ambiguïteit, twijfel, contrast	Idt
Verschillen(d)	Vragend, openbrekend	Op
Er zijn verschillende ...	Onderscheid in dienst van redenering	Idt & Re
Ander soort ...	Redenerend, functie in een analyse	Re
	Neutraal constaterend	-
Verschillen	In dienst van redenering	Re
Overeenkomsten & verschil	Gekoppeld aan een vraag	Op & Re
vergelijkingen	Gekoppeld aan eigen standpunt	Ep & Re
Anders (dan)	Constateren van ambiguïteit, twijfel, contrast	Idt
Dan is het anders	Constateren van onderscheid in dienst van redenering	Idt & Re
	In dienst van redenering of analyse	Re
	Vergelijkend	Re
	Neutraal constaterend	-
Een auto is wat anders dan een mens	Ambiguïteit exploiterend	Idt
Anders is het ...	In dienst van redenering	Re
Andersom	Mogelijkheid om het om te draaien, andersom te ...	Idt
omdraaien	Redenerend	Re
Eigenaardig	Ambiguïteit, 'nattigheid', 'het kan eigenlijk niet'	
Gek	Contra-intuïtieve gedachte	Idt
Raar (gevoel)	Waardeoordeel	Ep
Vreemd	Vragend	Op
'vaag', 'apart', 'stom'	'Ik vind het ...'	Idt & Ep
	Proberend 'een of ander raar iets'	Te
Ik snap/begrijp het niet	Verbaasd, zich verwonderend	Op & Ep
	Niet weten	Ep
Wat heb je eraan?	Verwondering	Op
	Suggestief, proberend	Te
	Afwijzend	-
Dat is wel zo	Zijn eigen mening wijzigend	Op
Dat wisten ze toch niet?	Verwonderend	Op
Maar ...	Twijfel, ambiguïteit,	
Tenminste	aan de ene kant, aan de andere kant	Idt
Maar niet helemaal ...	Tegenoverstellend	-
	Beperking	-
	Provocerend, actief onderscheid makend	Re
	Proberend	Te
	Stopwoord	-
Misschien (wel)	Twijfel, ambiguïteit	Idt
Mogelijk	Proberen, nieuwe suggestie, stel je voor	Te
	Inhoudelijke voorzichtigheid	Te
Misschien voelt een auto wel	Nieuwe suggestie, maar onzeker	Idt & Te
Misschien denkt die pen wel	Je weet niet of die pen denkt	Idt & Te

Hoewel, ofschoon	Twijfel, onzekerheid	Idt
Ondanks, tenzij	Inperkend	-
Ook al ...		
Terwijl		
Tenminste		
Kunnen	Capaciteit, vermogen, in staat zijn / gesteld worden	-
	'Het zit er (niet) in', mogelijkheid, suggestie	Te
	Ambiguïteit: het kan zus, het kan zo	Idt
Het zou kunnen	Ambiguïteit: het kan zus, het kan zo	Idt
Het zou moeten	Verwondering, wie weet?	Op
Zouden	Proberend	Te
Ik zou ...	Proberend	Te
	Oordeel, 'roletaking'	Ep
Het zou best kunnen	Proberend	Te
Er zijn best wel	Mogelijkheid	Te
Kan (toch) ook	Ambiguïteit, onbepaaldheid; kan zus, kan zo	Idt
Kan wel	'Oh ja ...', Verwondering	Op
Kan net zo goed	Suggestie, mogelijkheid	Te
Kan zus ... kan zo ...	Redenerend	Re
	Ambiguïteit in dienst van redenering	Idt & Re
	Persoonlijk of bepaald gebruik	-
	'Laat maar verder'	-
Kan wel ... maar toch ...	Ambiguïteit	Idt
Kan wel ... alleen ...	Ambiguïteit in dienst van redenering	Idt & Re
Dat kan toch niet (?)	Onzekerheid	Idt
Dat klopt toch niet (?)	Suggestie	Te
	Het kan niet anders	Re
	Proberend	Te
	Versterkend	-
	Intimiderend	-
	Openbrekend	Op
Kan niet anders	In dienst van redenering	Re
Van de vele mogelijkheden		
kan er maar een uitkomen		
Je kan nooit zeggen dat	Onzekerheid, ambiguïteit	Idt
bestaat niet	Retorisch middel om open te breken	Op
Ze zullen heus niet/wel	Proberend	Te
Niet zus, maar zo	Onderscheidend	Idt
Aan de ene kant ... aan	Ambiguïteit, grensgeval, relativiteit	Idt
de andere kant		
Of (disjunctie)	Twijfel, ambiguïteit	Idt
Of ... of ...	Tentatief, proberend	Te
Dit of dat, zus of zo	Onderscheid in dienst van redenering	Re
Tenzij	Opsomming, elkaar uitsluitend	-
Ondanks		

En (conjunctie)	Aan de ene kant en aan de andere kant	Idt
Ook/evengoed	Opsomming proberend	Te
Niet alleen	Opsomming (neutraal)	-
Dat heeft hij ook	Vergelijking in dienst van redenering	Re
'Niet alleen ... ook'	'Het eerste is te simpel, het is ingewikkelder'	Idt
Toch	Verwondering, is toch zo?	Op
Toch?	Het andere proberend	Idt
Kan toch ook	Proberend	Te
	Dat weet je toch? (verontwaardigd)	Te
	Verontwaardiging als stopwoord	-
Kan wel	Proberend	Te
	Zich ervan afmakend	-
Dat is toch ook/wel	pushend	Te
... ook nog	Ambiguïteit	Idt
Dat is toch	Informatie vragend	-
Juist	Ambiguïteit, tweespalt zaaiend	Idt
(Toch) juist wel / niet	Openbrekend	Op
	Treiterend	-
	Verbeterend	-
Denk ik	Meta-positie	Ep
Ik denk	Voorzichtig en afstand nemend	Te & Ep
Ik vind	Onzekerheid	Idt & Ep
	Proberend, voorzichtigheid	Te
	Stopwoord	-
... dat vind jij	Arbitraire karakter van oordeel	Idt
Niet kunnen weten	Scepsis,	
	Algemene vorm van niet-weten	Idt
Wij kennen ...	Wij i.h.b.	Ep
Wij kunnen waarnemen	Wij i.h.a.	-
Ik (wij) weet (weten) niet	Algemene vorm van twijfel, niet-weten	Idt & Ep
	Specifieke vorm, m.b.t. bepaalde situatie	Ep
Je weet het niet	Men weet niet, je kan het niet weten	Idt
	'Roletaking', ik kan me in het niet-weten verplaatsen	Ep
Ik denk, maar weet het niet		Idt & Ep
Ik zou wel eens willen weten/zien, hoe ...	Verwondering	Op & Ep
Weet ik veel	Met adequate betekenis	Ep
	Proberend	Te
	Zonder betekenis, stopwoord	-
Wie weet/zegt ...	Verwondering	Op
Als je het niet weet ...	Het zou kunnen	Idt
	Tentatief, proberend	Te
	Functie in redenering	Re
Gebruik van vergelijkende,	Te onderscheiden grensgeval	Idt
vergrotende of overtreffende	In dienst van redenering	Re
trap	Proberend	
	Versterking van oordeel	-

Goed, beter	Waardeoordeel	Ep
Het ligt eraan	Ambiguïteit: de ene keer wel, de andere keer niet	Idt
Het hangt ervan af	Twijfel, relativiteit, onbepaaldheid	Idt
Het heeft ermee te maken	Proberend	Te
	Bedenk wel, openbrekend	Op
	Een verband leggend, in dienst van redenering	Re
	Pushend, gevolgd door illustratie, voorwaardelijk	Te
	Zich ervan afmakend	-
	Beschrijvend: het ligt aan iets bepaalds	-
Dat ligt eraan wat je geest ervan vindt		Idt
Dat heeft er wel/niet(s) mee te maken	Vragend, op zoek naar relatie	Re
	Stellend, in dienst van redenering	Re
Dat maakt (niet) uit	Proberend	Te
	Beschrijvend, constaterend of zich ervan afmakend	-
Dat zegt niets	In dienst van redenering	Re
Vragen	Onderzoekend, divergent	Op
	Verwondering	Op
	Vragen om enkelvoudige informatie of uitleg	-
	Hoe bedoel je? (vraag naar info)	-
	Retorische vragen in dienst van redenering	Op & Re
	Suggestief, proberend (veelal met 'toch')	Te
	'Wat moet ik dan geloven? (suggestief, verontwaard.)	-
	Reflexieve vragen	Op & Ep
Ik vraag me af	Onderzoekend	Op & Ep
Dat is de vraag	Onderzoekend	Op
Wie zegt dat?	Provocerend, openbrekend	Op
	Spottend, pestend, niet serieus bedoeld	-
Oh ja	Toegeven, erkennen, 'je hebt gelijk'	Op
Dat is wel zo Daar heb je gelijk in Ik ben er anders over gaan denken Inderdaad	Belangeloze erkenning. 'niet eerder aan gedacht', bereidheid om ideeën te wijzigen niet afgedwongen omslag van mening	Op
	Anderszins	-
Hakkelende uitspraken	Proberend	Te
Stotteren / stopwoorden		-
Moeizame formuleringen	Poging om een gedachte te omschrijven	Te
Herhalingen en pleonasmen	Alleen indien ze proberend worden gebruikt	Te
	Bekrachtiging, vergroting van de verstaanbaarheid	-
Voorzichtigheiduitdrukking		Te
Onbedoelde tegenspraken	Zoekend, proberend	Te
Foutieve contaminaties	Zoekend, proberend, moeizaam formulerend	Te
Onconventioneel taalgebruik	Zoekend, proberend	Te
Nieuwe woorden	Zoekend	Te
Ongewone metaforen	Zoekend en vergelijkend	Te & Re

(Nou) Kijk ...	Proberend, zoekend	Te
... gewoon ...	Pushend	
Hoe zeg je dat	Moeilijk onder woorden te brengen	
Weet ik veel		
Ik bedoel ...		
... of zo(iets)		
... zeg maar	Stopwoord	-
Op een of andere manier		
Waarschijnlijk	Proberend	Te
Als ...	Contrafactische hypothese, stel dat, stel je voor ...	Te
(gevolgd door suggestie)	Voorwaardelijk	-
Als ... (dan ...)	Contrafactische hypothese + gevolgtrekking	Te & Re
Stel (dat) ...	Voorwaardelijk in dienst van redenering	Re
	Beschrijving van chronologische volgorde, constatering	-
	Gevolgd door vraag	Op & Te
	Gevolgd door ... (dan) vind ik	Op & Ep
Als je ..., dan denk ik dat ...	Meta-opmerking	Ep
Bijvoorbeeld	Beschrijving van een anekdote	An
Een voorbeeld	Proberend, hypothetisch	Te
'Kijk ...'	Stopwoord	-
Zich realiseren		Ep
Zich voorstellen		
Reflexieve werkwoorden	Met cognitieve betekenis	Ep
	Zonder cognitieve betekenis	-
Cognitieve werkwoorden in	Meta-opmerking	Ep
de 'ik' -vorm:	Als voorbode van uitleg of directe rede	-
Ik geloof, twijfel, denk,	Proberend, moeizaam formulerend	Te
vind (niet), overweeg,	Als stopwoord	-
beschouw, weet (niet), wil,	Expliciete onzekerheid	Idt
verwacht, ervaar ...		
Ik bedoel	Met adequate betekenis	Ep
	Proberend, voorzichtigheid	Te
	Stopwoord	-
Ik heb wel eens ervaren/	Gevolgd door anekdote	An
gedroomd		
Volgens mij/ons	Meta-opmerking	Ep
Voor mij/ons	Stopwoord	-
Waardeoordelen	Waardeoordeel van de spreker, niet instrumenteel	Ep
	'Goed omdat': Instrumentele middel -doel -relatie	Re
	Algemeen voorschrift	-
	In combinatie met 'eigenlijk'	Idt
Ze zeggen/denken	Indien spreker zich in die rol verplaatst	Ep
	Algemene constatering	-
We hebben het erover	Meta-opmerking	Ep

Moeten	Voorschrift, verplicht zijn/voelen	-
Hoeven	Willen	-
Het zou moeten	Noodzakelijk gevolg, het kan niet anders	Re
	Aannemelijk zijn	Re
	Proberend, tentatief	Te
Het heeft nut ...	Praktische redenering	Re
Dat moet je zien als ...	Soort verplichting	-
Het hoeft niet (per se)	Ambiguïteit, nattigheid	Idt
	Openbrekend, corroborerend	Op
	Tentatief, proberend	Te
	In dienst van redenering of analyse	Re
	Herhalend, bevestigend	-
Het ligt eraan, het hoeft niet	Relativerend	Idt
Mogen	Persoonlijk waardeoordeel	Ep
	Algemeen waardeoordeel, morele verplichting	-
... dan ...	Gevolgtrekking	Re
Relatie aangevende	Causaal verband	Re
voegwoorden:	Chronologisch verband	-
Omdat, want, indien,	Voorschrift	-
doordat, dus ...	Uitleg, betekenisgevend	-
	Stopwoord	-
Dat komt doordat ...	In dienst van redenering	Re
Zo ...	En zo komt het dat	Re
Omkeringen	In een redenering	Re
Afleidingen	Deductief redeneren	Re
	Inductief redeneren	Re
(Net) Zoals ...	Aangeven van overeenkomsten en verschillen in dienst	Re (&
(gevolgd door illustratie)	van redenering	Te/An)
Hetzelfde als ...	Associatieve situatiebeschrijving	Te
	Associatieve anekdotebeschrijving	An
	Definitie, constatering	-
Betekenen	In dienst van redenering	Re
Dat zou betekenen	Uitleg, definitie	-
Dat betekent dat ...		
Dat is niet eerlijk	Ambiguïteit, nattigheid	Idt
	Niet volgens de wet / de regels	-
	Waardeoordeel	Ep
Dat is niet eerlijk, want ...		Re
Het kan gelogen zijn	Je weet het niet, is ook mogelijk	Idt
Iemand kan liegen		
Natuurlijk ... (gevolgd door		Re
verklaring/reden)		
Citaat	Verwijzend naar een specifieke gebeurtenis/situatie	An
	Inhoud van citaat	-
	Open vragen	Op
	Ironische opmerkingen	Idt

Idt: Indecisive thinking Te: Tentative behaviour Re: Reasoning quality
Op: Openness Ep: Epistemic position An: Anecdotal quality

Summary

Why do children and adolescents sometimes perform such authentic thinking patterns when dealing with philosophical questions? Is it based on a stable philosophical quality? Is it linked to personality traits, in other words: a (philosophical) talent? And if so, how could that be determined scientifically? Intrigued by these questions, a systematic study on the philosophical quality of such thinking patterns among 10 - 20 years old was undertaken. Youngsters from that age-range were selected because their trains of thought are considered to be more genuine and relatively unaffected by learned reaction patterns, knowledge and prejudices accumulated over time. This thesis aims to establish whether philosophical talent exists, and if so, to determine the nature of such potential talent in youngsters. This was achieved by addressing a special quality of philosophical thinking patterns through almost 100 standardised discussions with youngsters. If detected, such a quality might be given a position in a nomological network of relations among other individual and group characteristics, and be potentially labelled as a talent. This thesis consists of nine chapters, beginning with a general introduction to the theme in Chapter 1.

'Philosophical quality' strongly relates to concepts of 'philosophy', 'philosophising', and 'wisdom'. To describe 'philosophical quality', a common understanding of these terms would be needed. To that end, in Chapter 2 it is shown that three main features of philosophy were derived from history and from descriptions by leading current philosophers about the mental activity of philosophising: 1) analysing and reasoning qualities; 2) qualities detecting ambiguities, vagueness, uncertainty and borderline explorations; and 3) qualities of moving smoothly from theory to practice and vice versa, realised within a wide framework with connections between knowledge and experience. Empirical studies on wisdom in the 1980s and 1990s uncovered similar individual qualities. By then, attempts were being made to correlate these tendencies with measurable data derived from individuals, such as intelligence, age or personality traits. Two levels of conceptual estimation of thematic developments and of empirical investigation of individual performances must be brought together in a wider scheme.

It is now accepted philosophical quality has at least three of philosophy's traditional main features in philosophy in common; conceptual considerations derived from these are matched with observations of real life expressions in Chapter 3. Such expressions are uttered by youngsters while discussing philosophical topics designed to detect philosophical qualities of thinking patterns both at individual and at group level. Their mental explorations of philosophical topics cover a wide spectrum from tentative behaviour, trying to 'capture' the unknown and proceeding along their own and non-predictable way of thinking. This seemingly chaotic exploration result hampers judgement by conventional criteria, only referring to discussion content or to evaluative norms like presuppositions of maturity, requirements of correctness, and justified cultural agreements. Vagueness and ambiguity, as the pre-eminent drivers in the process of evolving philosophically qualified concepts and ideas, are greatly welcomed. Here, these are treated as observable and

countable occurrences. By evaluating adolescent thinking patterns as a collection of such observable occurrences potential discrepancies between theoretical and empirical psychological approaches are bridged.

At this stage in the study, a conceptual framework and a measuring instrument are needed to uncover the philosophical quality of thinking patterns. These are provided in Chapter 4, in which is described the development of such a framework and the empirical way in which oral utterances by youngsters are placed in a standardised format and systematically addressed and measured. The conceptual framework consists of five indicators, distinctly assessing aspects of the philosophical quality of thinking patterns through oral expressions: Indecisive thinking (Idt), Openness (Op), Tentative behaviour (Te), Epistemic position (Ep), and Reasoning quality (Re), to some degree supplemented with a sixth indicator, Anecdotal quality (An).

Occurrences of these indicators in youngsters' expressions in dedicated philosophical discussions must now be registered and counted. To this end, a specific instrument has been developed for this study, namely the tetralogue.

Tetralogues can measure indicator frequencies (scores) of utterances and mutually compare these over time. A tetralogue is a standardised philosophical discussion in which four participants exchange their trains of thought as they address a philosophical topic. Tetralogues are ignited by key questions, and are selected by participants on a voluntary basis. Tetralogues are chaired by qualified experts. These chairpersons are bound by predetermined rules concerning interference. In total, 95 tetralogues are recorded (video- and audio taped) with 302 participants in three age categories: 11 to 13, 14 to 16, 17 to 18 years old. Only the 11 to 13 and 14 to 16 year old categories (70 tetralogues with 215 participants) were available for analysis in this thesis. Two educational levels for participants were registered: high – VWO-HAVO; and low – VMBO. Furthermore, two levels of life experience were distinguished: youngsters with a regular life course, and those with an irregular life course. When formatted, the 70 tetralogues revealed a total of 14,393 utterances. An utterance consists of one to about ten sentences, unbrokenly expressed by one participant until interrupted by a co-participant. All utterances were transcribed, formatted and checked for indicators of philosophical quality. These formats permitted philosophically qualified thinking patterns to be measured empirically and processed quantitatively. The objectivity and reliability of this scoring method has been demonstrated by proving an inter-rater agreement, internal consistency of indicator frequencies, and by good split-half and test-retest coefficients. With a homogeneity alpha of 0.8, the assessed indicators represent the measured quality of utterances by youngsters in tetralogues.

Next, two numerical indices were constructed: pq and PQ. Respectively, these indices reflect philosophical quality of individually and collectively performed thinking patterns in tetralogues. The pq index (individual performance) is based on a balanced ratio of indicator frequencies in line with the main features of philosophy. The PQ index (group performance) was developed to measure the philosophical power of joint performance exceeding that of individual contributions. Dialogue events of jointly generated, qualified combinations of indicator frequencies have

also been taken into account. These steps produced evidence that places philosophical quality on record according to the two types of indices. Finally, the main aim of this study was met by successfully corroborating the construct validity of the tetralogue, by comparing calculated indices with previous estimates of philosophical quality, and by controlling the assumed relations of the nomological network. The last step is set out in Chapter 6.

Once the very existence of a philosophical quality had been assessed and quantified on an individual and group level, the philosophical topics discussed in tetralogues were assigned to classic philosophical categories in Chapter 5. These categories include metaphysics and epistemology, anthropology, ethics, and topics dealing with meaning and demarcation problems. Next, outcomes of individually (pq) or group (PQ) performed philosophical qualities were checked against these categories. As different themes do not lead to significant differences in pq and PQ indices and as all categories of thematically grouped tetralogues show similar patterns of indicator frequencies, it was concluded that the philosophical qualifications of tetralogues are unrelated to the four philosophical categories. This demonstrates the general character of the conceptual framework and of the tetralogue as an instrument of philosophical enquiry irrespective of its theme. Nevertheless, differences can be noticed between participant preferences for key questions and themes: anthropological key questions seem to be more popular among females than among males, and more among the low educated than among highly educated youngsters. Female groups and higher educated participants preferred ethical themes.

As it is found that philosophical quality is basically unrelated to philosophical themes, and that tetralogues are instruments of a general nature, it became possible to check if philosophical quality is perhaps related to specific participant or group characteristics. This is considered in Chapter 6, where it is demonstrated that philosophical quality is significantly and positively correlated with *openness to experience* as a personality trait, with educational level, and to intelligence to a limited extent. The construct validity of the tetralogue as a measuring tool for philosophical quality is further corroborated by the fact that nomologically expected relations are replicated empirically. So, PQ and pq prove to be valid indices capable of measuring the philosophical quality of thinking patterns.

However, the relation between philosophical quality and individual and group characteristics appears to be rather complicated. No such relation is found between philosophical quality (pq and PQ) and age, although age variance was affected by the transition from primary to secondary school. Relations between philosophical quality and measures for intelligence appear to be significant but low, indicating a difference between philosophical qualified and convergent thinking patterns. They appear to be affected by the threshold values and meaning of Grade Point Average (GPA) for language performance. Presupposed relations between individual philosophical quality and irregular life course were not found, nor between PQ indices and heterogeneity in life courses among tetralogue participants. This was (at least partly) attributed to the apparent impossibility of determining the characteristics of a 'regular' life course. At the same time, educational level seems to affect the relationship between philosophical quality and life course.

Despite this, the question arises as to whether PQ indices might be dependant on the degree of intervention and management style of a chairperson? This is explored in Chapter 7. Design and realisation of tetralogues were supposed to vary in degree of intervention and management styles. This study shows no significant differences in PQ in relation to management style or degree of intervention. However, if a tetralogue chair has less philosophical experience, indicator frequencies for epistemic position and for anecdotal quality appear to be slightly higher, while indicator frequencies for openness seem to decrease with more than average chair interruptions. These outcomes suggest a qualitative relation with management styles dependent on the chair's philosophical experience rather than with quantitative differences in PQ indices.

In order to check whether philosophical quality changes over time in a transition period from primary to secondary school, a limited follow-up study, lasting two years and four months, was undertaken with four boys aged between 11 and 12 years at the onset of this study. During this period, the boys moved from primary school to a secondary school of low-educational level. Over time, no linear change in philosophically qualified thinking patterns could be determined with respect to individual performances, group performances and indicator frequencies. However, if changes in pq indices were described according to a cubic model, trends in pq indices became evident: these increased during the boys' last months in primary school and continued over the next few months. A turning point can be recognised some months after departure from primary school, with downward trends appearing on statistical grounds, biographically through school transition, and circumstantially through environmental events. Such events become evident during tetralogues and are revealed through disruptive behaviour, neglect of appointments, and in being denounced. These circumstances mirror the influence of school environment and peers. At the same time, a decrease in references to real life evidence is reflected in a significant decrease in anecdotal quality scored in utterances. The results of this follow-up study, described in Chapter 8, do not contradict the idea of philosophical quality in terms of talent.

In the last chapter (Chapter 9), all the findings of this research are placed in perspective of the thesis' aims and of some traditions or discourses. The main aim of this study is met by establishing that objective and reliable assessment of philosophical quality provides a valid measure of philosophically qualified thinking patterns. This was accomplished using by two types of indices: pq for individual performances and PQ for group performances. By successfully quantifying this quality, the research has specified generalisations of philosophy and clarified philosophy as mental activity; it has also confirmed that philosophy is a distinguished domain of cognitive behaviour. Philosophically qualified thinking patterns are linked to rationality as a faculty of mind, including non-rational patterns such as intuition, imagination and emotional considerations on the condition that these types of expressed thinking can be approached rationally.

As philosophical quality is determined among young people, the results of this study may be compared to studies and comments on philosophy with children. However, most of the observations in the literature address educational goals and

the salutary effects of the *Philosophy for Children* programme. Some similarities are noticed between the approaches of Matthews (1980, 1994, 1998), Freese (1990) and this research, but only the latter has provided a measurement that determines and compares different philosophical qualities. Empirically obtained results of philosophical quality resemble those of wisdom. Both indices are related to educational level and openness to experience as a personality trait, while no relationship with age could be assessed. In terms of cognitive development, philosophically qualified thinking patterns do not converge and consequently are not related to traditional measurements of cognitive and moral development. Nevertheless, the results of this thesis contribute to a deeper understanding and to potential measurement of sensitivity to pre-conceptual knowledge.

Using the term *philosophical talent* as an equivalent of philosophical quality is justified in the present research by demonstrating a replication of talent features according to the model of Ziegler and Heller (2000). This is correct since the established philosophical quality shows individual stability, independence of age, relation to openness to experience as a personality trait, rooting in rational beliefs, and the possibility of exploitation in cooperation with the environment. Finally, the applicability of pq and PQ indices were examined with respect to philosophical education and evaluation, to selection of personnel, to developing a critical attitude, and to a disinterested drive into scientific research.

Samenvatting

Waarom komen kinderen en adolescenten soms met zulke authentieke gedachtegangen wanneer ze geconfronteerd worden met filosofische vragen? Berust dit verschijnsel op een bepaalde, gegeven filosofische kwaliteit? Bestaat er een relatie met een persoonlijkheidskenmerk, met andere woorden: met een filosofisch talent? En als dat zo is, hoe zou dit dan wetenschappelijk kunnen worden vastgesteld? Het zijn deze vragen die hebben geleid tot een systematisch onderzoek naar de filosofische kwaliteit van het denken van jongeren tussen de 10 en 20 jaar. In deze leeftijdsgroep zijn gedachtegangen over het algemeen nog nauwelijks beïnvloed door een overdaad aan kennis, cultureel gevormde reactiepatronen, clichés en vooroordelen en kunnen daarom veelal als origineel worden beschouwd. Dit onderzoek beoogt het bestaan van een filosofische kwaliteit aan te tonen en vast te stellen of daarbij al dan niet sprake is van een talent. Daartoe zijn bijna 100 gestandaardiseerde filosofische gesprekken met jongeren uitgevoerd. Alle uitspraken daarin zijn getest op filosofische kwaliteit met behulp van indicatoren die op hun beurt weer afgeleid zijn uit de filosofische literatuur. In negen hoofdstukken worden de resultaten van dit onderzoek samengevat; hoofdstuk 1 omvat de algemene inleiding op deze studie.

Om een 'filosofische kwaliteit' te kunnen beschrijven, wordt eerst gekeken wat 'filosofie', 'filosoferen' en 'wijsheid' met elkaar gemeen hebben. In hoofdstuk 2 zijn drie hoofdkenmerken van het filosoferen geformuleerd, afgeleid uit historische ontwikkelingen in de filosofie en uit beschrijvingen van hedendaagse filosofen die zich expliciet hebben uitgelaten over de activiteit van het filosoferen. Deze conceptuele kenmerken zijn 1) een analytische en redeneerkwaliteit; 2) een gevoel voor ambiguïteit, vaagheid, onzekerheid en voor het aftasten van grenzen; 3) het flexibel heen en weer kunnen springen van theorie naar praktijk en omgekeerd en dat binnen een geïntegreerd raamwerk van kennis en ervaring. Empirisch wijsheidsonderzoek in de jaren tachtig en negentig van de 20ste eeuw legt soortgelijke kenmerken bloot en probeert deze te correleren met meetbare persoonsgegevens, zoals intelligentie, leeftijd en persoonlijkheidskenmerken. De conceptuele en de empirische niveaus waarop filosofische kwaliteit benaderd kan worden, dienen nu bijeen gebracht te worden.

Deze hoofdkenmerken van filosoferen worden in hoofdstuk 3 getoetst in actuele gesprekken waarin adolescenten filosofische problemen onderzoeken. Zij gebruiken daarin uitdrukkingen waarin ze een filosofische kwaliteit van hun denken aan de dag leggen, zowel op individueel als op collectief niveau. Deze uitdrukkingsvormen laten een breed spectrum zien van tentatief gedrag, pogingen om het onbekende in woorden te vangen, en volgen een volstrekt eigen, onvoorspelbare lijn. De soms schijnbaar chaotische resultaten van zulk denken kunnen maar voor een deel met conventionele criteria worden beoordeeld. Dat komt omdat zulke criteria vooral verwijzen naar de inhoud van het gesprek en gebruik maken van evaluatieve normen, als rijpheid, correctheid en cultureel gerechtvaardigde afspraken. Daarentegen vormen vaagheid en ambiguïteit juist bij uitstek een bron waaruit filosofisch gekwalificeerde gedachtegangen ontwikkeld kunnen worden.

Opmerkingen die op vaagheid, ambiguïteit en onzekerheid wijzen, moeten daarom worden waargenomen en uiteindelijk geteld. Door nu het denken van adolescenten als een verzameling waarneembare gebeurtenissen op te vatten, kan een mogelijk gat tussen een theoretische en empirische benadering worden overbrugd.

Voor het systematisch vastleggen van een filosofische kwaliteit in het denken van adolescenten is een conceptueel kader vereist en een meetinstrument. In hoofdstuk 4 worden deze ontwikkeld en beschreven. Het conceptuele kader bestaat uit vijf indicatoren die ieder afzonderlijk een of meerdere aspecten van de filosofische kwaliteit in de geuite gedachten vastleggen. Het zijn: Onbepaald denken (Idt), Openheid (Op), Tentatief gedrag (Te), Epistemische positie (Ep), en Redeneer kwaliteit (Re). Dit vijftal wordt onder bepaalde omstandigheden aangevuld met een zesde indicator: Anekdotische kwaliteit (An).

Deze indicatoren zijn herkenbaar in mondelinge uitlatingen van jongeren in gesprekken over filosofische thema's. Zulke uitlatingen worden geregistreerd en zodanig geformatteerd dat ze geschikt worden voor telling en systematische analyse. Dit gebeurt via een speciaal daartoe ontwikkeld instrument: de tetraloog. Tetralogen zijn ontworpen om het voorkomen van de indicatoren in uitdrukkingen te meten (scores) en om het optreden van indicatorcombinaties bij verschillende individuen op verschillende tijdstippen met elkaar te kunnen vergelijken. Een tetraloog is een gestandaardiseerd filosofisch gesprek waarin vier adolescenten (deelnemers) gedachten met elkaar uitwisselen naar aanleiding van een filosofisch probleem dat door de deelnemers vooraf zelf wordt aangedragen. Tetralogen worden begeleid door een filosofisch of psychologisch geschoold persoon die het gesprek alleen mag onderbreken op tevoren vastgestelde gronden. In totaal zijn 95 tetralogen op video- en audioband vastgelegd, met 302 deelnemers in drie leeftijdscategorieën: 11 tot 13, 14 tot 16 en 17 tot 18 jaar. Zeventig daarvan met 11 tot 13 jarigen en met 14 tot 16 jarigen (215 deelnemers) zijn in het kader van dit onderzoek uitgewerkt. Deelnemers zijn ingedeeld in twee opleidingsniveaus: hoog: HAVO-VWO en laag: VMBO, en daarnaast in twee niveaus van levenslopen: jongeren met een reguliere en met een irreguliere levensloop. De 70 geanalyseerde tetralogen brachten een totaal van 14.393 zegbeurten voort. Een zegbeurt bestaat uit één tot tien ononderbroken zinnen, uitgesproken door eenzelfde deelnemer totdat deze wordt onderbroken door een ander. Alle zegbeurten zijn op schrift gesteld, geformatteerd en gecheckt op het voorkomen van indicatoren voor filosofische kwaliteit. Door deze procedure kan filosofisch gekwalificeerd denken empirisch gemeten worden en gekwantificeerd. Objectiviteit en betrouwbaarheid van deze scoringsmethode zijn aangetoond met een inter-raterovereenstemming, interne consistentie van indicatorvoorkomens, en door goede splithalf- en test-retest-coëfficiënten. Met een homogeniteitindex alpha van 0,8 blijken de vastgestelde indicatoren inderdaad de filosofische kwaliteit in de uitlatingen van jongeren gedurende een tetraloog te representeren.

Vervolgens zijn twee numerieke indices ontworpen: pq en PQ, om de filosofische kwaliteit van respectievelijk individueel en collectief voortgebrachte gedachtegangen in een tetraloog weer te geven. De pq-index (individueel gepresteerd) berust op een verhouding tussen de vijf indicatorvoorkomens gemodelleerd naar de

basiskenmerken van het filosoferen. De PQ-index (groepsprestatie) is ontwikkeld om de filosofische kwaliteit van het groepsoptreden te meten die de som van individuele bijdragen te boven gaat. De aard van de samenspraak in een tetraloog met de daarin voorkomende combinaties van indicatoren moet in de PQ-index ook tot zijn recht komen. Beide indices maken het mogelijk de filosofische kwaliteit van een individueel en van een groepsoptreden in een getal uit te drukken. Aan de validiteit van de gemeten filosofische kwaliteiten wordt voldaan door deze te vergelijken met andere, onafhankelijke schattingen van de betreffende filosofische kwaliteiten. De berekende PQ-indices bleken significant te correleren met in een vroeg stadium gedane grove schattingen op een 5-puntsschaal. Individuele pq-indices bleken eveneens significant te correleren met pq-indices van dezelfde individuen in andere tetralogen. Verder is de validiteit vastgesteld door na te gaan of relaties die op theoretische gronden zijn aangenomen (nomologisch netwerk) ook proefondervindelijk worden teruggevonden. De resultaten daarvan worden in hoofdstuk 6 beschreven.

In hoofdstuk 5 wordt eerst nagegaan of de gemeten filosofische kwaliteit op individueel en op groepsniveau mogelijk afhangt van de gebruikte filosofische vraagstelling. Hiertoe zijn alle beginvragen en uitgewerkte filosofische thema's van tetralogen gecategoriseerd volgens een klassieke indeling: metafysica en epistemologie, antropologie, ethiek, en onderwerpen met betrekking tot betekenis- en demarcatieproblemen. Individueel, noch collectief totstandgebrachte filosofische kwaliteiten (pq en PQ indices) bleken met deze categorieën te correleren. Bovendien openbaarde zich bij alle filosofische thema's eenzelfde patroon van indicatorvoorkomens. Het conceptuele raamwerk van indicatoren en het instrument van de tetraloog om filosofische kwaliteit te meten, blijken dus onafhankelijk van het filosofische thema te werken. Er kunnen echter wel voorkeursverschillen voor filosofische thema's optreden tussen verschillende groepen deelnemers. Antropologische vragen blijken eerder gekozen te worden door groepen met overheersend meisjes en lager opgeleiden. Meisjes en hoger opgeleiden prefereren eerder ethische thema's.

Nu blijkt dat de filosofische kwaliteit van een gesprek onafhankelijk is van de (filosofische) vraagstelling kan onderzocht worden of er mogelijk relaties bestaan tussen de gemeten resultaten (pq en PQ) en specifieke kenmerken van individuen of groepen. Dat gebeurt in hoofdstuk 6. Individuele pq-indices blijken significant te correleren met het persoonlijkheidskenmerk *openness to experience*, met het individuele opleidingsniveau, en – ofschoon laag en minder significant – ook met intelligentie. Dat verwachte relaties in de meetresultaten teruggevonden worden, ondersteunt de validiteit van de tetraloog als meetinstrument voor filosofische kwaliteit. PQ en pq zijn dus geldige indices om de filosofische kwaliteit van het denken van jongeren te meten.

Maar het verband tussen filosofische kwaliteit en individuele of collectieve eigenschappen is vaak gecompliceerd. Hoewel er geen relaties tussen pq- of PQ-indices en leeftijd gevonden zijn, wijken de meetresultaten van basisschoolleerlingen op enkele punten af van die van middelbare scholieren. De zwakke correlatie tussen filosofische kwaliteit en intelligentie wijst erop dat er ook een verschil is tussen

filosofisch denken en convergent denken. Er lijkt een drempelwaarde te bestaan voor verbale intelligentie, gemeten met het gemiddelde schoolcijfer voor talen. Boven deze waarde verdwijnt de relatie tussen filosofische kwaliteit en opleidingsniveau. Tenslotte is een veronderstelde relatie tussen individuele filosofische kwaliteit en irreguliere levensloop niet in de meetresultaten teruggevonden. Dit geldt evenmin voor die tussen PQ-indices en heterogeen samengestelde groepen (irreguliere of reguliere plus irreguliere levenslopen). De reden hiervoor is wellicht dat het zeer moeilijk is vast te stellen wat een *reguliere levensloop* precies inhoudt. De relatie tussen filosofische kwaliteit en levensloop blijkt bovendien beïnvloed door het opleidingsniveau.

In hoofdstuk 7 wordt onderzocht of PQ-indices mogelijk beïnvloed kunnen zijn door aard of aantal van de interventies door de gespreksleider. Tetralogen zijn ontworpen zonder vaste omschrijving van het type gespreksleider en vertonen verschillen in het aantal interventies. Dit onderzoek toont aan dat er geen significante (kwantitatieve) verschillen in filosofische kwaliteit bestaan tussen verschillende leidersstijlen en tussen het aantal interventies. Wel zijn er kwalitatieve verschillen geconstateerd in leidersstijlen die uitgedrukt kunnen worden in specifieke indicatorvoorkomens. Wanneer een tetraloog door een filosofisch minder gekwalificeerde wordt geleid, nemen de indicatorvoorkomens van epistemische positie en anekdotische kwaliteit een weinig toe. Bij bovengemiddeld interrumperen door de gespreksleider neemt de openheid (vragen) af.

In hoofdstuk 8 wordt onderzocht of zich veranderingen in filosofische kwaliteit voordoen binnen één individu en binnen één groep gedurende de overgang van basisschool naar middelbare school. In een beperkte follow-upstudie worden vier jongens twee jaar en vier maanden lang gevolgd. Aan het begin van deze periode zijn ze 11 of 12 jaar oud en zitten in groep 8 van de basisschool. Daarna gaan ze naar verschillende VMBO-scholen (met een relatief laag opleidingsniveau). In deze periode is geen lineaire verandering waargenomen in het filosofisch denken van de jongens, noch op individueel, noch op groepsniveau. Wanneer de veranderingen in pq indices echter worden beschreven volgens een kubisch model worden wel enige trends geconstateerd: een stijging van de filosofische kwaliteit tijdens de laatste maanden op de basisschool en de eerste maanden op de middelbare school, gevolgd door een omslagpunt, waarna de filosofische kwaliteit afneemt. Deze afname manifesteert zich in 3 opzichten: statistisch, biografisch (verhuizing naar een lager soort opleiding) en verstorende omstandigheden, zoals het vaak voorkomen van irrelevante en ondermijnende opmerkingen, afspraken die niet worden nagekomen en een aanhoudend negeren van één van de jongens. Deze omstandigheden weerspiegelen mogelijk de invloed van leeromgeving en vrienden. Tegelijkertijd wordt er in de gesprekken steeds minder verwezen naar concrete gebeurtenissen uit het dagelijkse leven en neemt de anekdotische kwaliteit in zegbeurten significant af. Deze resultaten corresponderen met het denken over filosofische kwaliteit in termen van talent.

Alle bevindingen van dit onderzoek worden in hoofdstuk 9 vergeleken met de doelstelling en met enkele tradities en in het denken over denken. Aan de belangrijkste doelstelling van het onderzoek is voldaan: het objectief, betrouwbaar

en valide vaststellen van het bestaan van een filosofische kwaliteit als een meetbare maat voor filosofisch gekwalificeerde gedachtegangen. Daarmee verheldert en specificeert dit onderzoek vage noties van filosoferen als mentale activiteit. Ook wordt hierdoor de notie van filosofie als specifiek kennisdomein versterkt.

Omdat het vaststellen van een filosofische kwaliteit is gebeurd bij jongeren ligt het voor de hand de onderzoekresultaten te vergelijken met die van het filosoferen met kinderen en met verschillende commentaren daarop. De meeste commentaren betreffen echter educatieve doelen van en propaganda voor het programma *filosoferen met kinderen*. Benaderingen van Matthews (1980, 1994, 1998) en Freese (1990) lijken nog het meest op die van dit onderzoek, behalve dat die geen meetinstrument aanreiken om filosofische kwaliteiten vast te stellen en te vergelijken.

Empirisch verkregen resultaten van filosofische kwaliteit kunnen vergeleken worden met die van wijsheid. Indices voor pq en indices voor wijsheid zijn beide gerelateerd aan opleidingsniveau en aan *openness to experience* als persoonlijkheidskenmerk; een significante correlatie met leeftijd is niet gevonden. Met betrekking tot cognitieve ontwikkeling dient benadrukt te worden dat filosofisch gekwalificeerde gedachtegangen niet convergeren en dus ook niet gemeten kunnen worden met traditionele maten voor cognitieve en morele ontwikkeling. De resultaten van dit onderzoek scheppen echter wel meer inzicht in de gevoeligheid voor preconceptuele kennis.

Het gebruik van het begrip *Philosophical Talent* (filosofisch talent) voor de gevonden filosofische kwaliteit wordt gerechtvaardigd door de aangetoonde overeenkomst tussen eigenschappen van pq en eigenschappen van een talent volgens het Ziegler en Heller – model (2000): de vastgestelde filosofische kwaliteit laat een individuele stabiliteit zien, is onafhankelijk van leeftijd, verbonden met *openness to experience* als persoonlijkheidskenmerk, wortelt in een rationele overtuiging en kan geëxploiteerd worden in samenwerking met omgevingsfactoren.

Tenslotte is gekeken naar toepassingsmogelijkheden van indices voor pq en PQ in het filosofieonderwijs en in de beoordeling van filosofische prestaties, in selectieprocedures voor bepaalde functies en beroepen, ten behoeve van het ontwikkelen van een kritische houding en van een belangenloze inspiratiebron in wetenschappelijk onderzoek.

Cooperating schools and institutes

Alphen aan de Rijn:	RIAGG, hulpverlening jongeren
Amersfoort:	Johan van Oldenbarneveldt Gymnasium
	ROC de Amerlanden
	Individual adopted adolescents
Amsterdam:	't Koggeschip
Bladel:	Pius X Scholengemeenschap
Brussels (Belgium):	Centrum Deeltijds Leren
	Wittouckschool
Den Haag:	J. C. Pleysierschool
Dinxperlo:	SG Schaersvoorde
Evergem (Belgium):	Het Molenschip, internaat voor schippers- en kermiskinderen
Haarlem:	Damiate College
	Schoterkring
	Stedelijk Gymnasium
Hengelo:	Twickel College
Middelburg:	Christelijke Scholengemeenschap Walcheren
Nijmegen:	De Monnikskap
Sassenheim:	Teylingereind (gesloten justitiële jeugdinrichting)
Utrecht:	Wittevrouwen Jenaplan-basisschool
Zeist:	Bartiméus Onderwijsinstelling (onderwijsbegeleiding slechtzienden)
	Jordan College

Acknowledgement

I am much indebted to my two promoters, Peter Heymans and Karel van der Leeuw: to Peter for his ability and readiness to explore new ventures in science and to Karel for his erudite comments, for his scientific guidance and inspiration through almost 25 years of my professional carrier.

I am most thankful to all 281 youngsters for their lively discussions and genuine thoughts in the numerous tetralogues we shared. Their preparedness and eagerness to participate in these sessions provided me with an invaluable dataset and inspired me to proceed with this activity, so essential for success in this endeavour. I would like to express my sincere gratitude to the leaderships of selected schools and institutes for their fine cooperation. I would like to highlight four of them in particular: Jantje Jurgens, (Schoterkring, Haarlem) for selecting and facilitating the follow-up group of four boys, Ton Jansen (Stedelijk Gymnasium, Haarlem) for disclosing one of the most renowned categorical grammar schools of The Netherlands, Ton Sernee (Damiate College, Haarlem) for his repeated assistance in approaching a substantial group of seemingly indifferent VMBO students, and Tom van Riel (Monnikskap, Nijmegen) for accessing a population of physically disabled participants. His constructive comments on the main chapter of this thesis are highly appreciated.

During my research I felt most welcome in the capacity group developmental psychology. Here I received feedback by many scholars and scientists, in particular Jan ter Laak, who has always been extremely supportive to assist this philosopher during her first steps in the field of methodology. I am very grateful to Gerard Brugman for inspiring me with wisdom, for his role as experienced co-judge, but most of all for introducing me in the world of Shostakowitsj. From my friends outside this scientific bastion in Utrecht I would like to express my sincere gratitude to Henk Becker, who has acted as impartial co-judge for several tetralogues and who was been so kind to play a major role in editing the references of this study. I am most indebted to Johannes Netelenbos for his essential contributions to data importing of 14,393 utterances in SPSS files, providing me with some critical comments.

Without the great support of David Huntley this thesis would have remained unreadable. I thank him very much for his corrections and suggentions to improve the English. My friend Edward Derbyshire completed this process for some chapters. I also thank Wendy Davies for translating the real life discussions with children in Chapter 3.

Through this thesis I have attempted to construct a metaphorical bridge between earth and heaven. This was done not only with respect to tangible – measurable – and conceptual affairs but also proved to operate as a mediator between geology and philosophy. Philosophers may learn a lot of practical (earthly) matters from geologists. Above all, I am deeply grateful to Eduardo de Mulder for his continuous mental encouragement and support. In particular, I thank him for introducing me

into the basics of scientific writings and for his critical comments. All mentioned persons dedicated to the project and rely in good faith on its performer.

Finally, I thank *Beethoven*, *Shostakowitsj*, and *Mahler* for their patient and unquestioned company during all those long car trips between Haarlem and Utrecht. They provided me with just that extra shot of adrenaline to begin my days in research and with the necessary relief to recover my brains for the next day.

Curriculum vitae

Thecla Rondhuis, born in 1950, graduated from high school in 1969 (HBS-B). She studied biology at Utrecht University until 1973. After a 10-years break while raising four children, she commenced her studies philosophy in the University of Amsterdam, which were successfully completed in 1990. From 1983 onwards, Thecla Rondhuis chaired philosophical sessions with children at primary schools on a weekly basis. In 1987 she co-founded the Centre of Philosophy for Children in The Netherlands. Thecla Rondhuis produced in 1993 and 1994 a series of TV programmes where groups of children discussed philosophical topics. Her first book: *Filosoferen met kinderen* (*Philosophising with children*) was published in 1994. From 1996 till 2002, she wrote monthly columns on children's authentic trains of thought in *Filosofie Magazine*, and in several other magazines. In 1997, she started her PhD research on philosophically qualified thinking patterns of children and youngsters at Utrecht University. Three years later, in 2000, she started teaching philosophy at De Monnikskap, a VWO school for physically handicapped adolescents. In 2001, she published a second book on philosophising with children: *Jong en Wijs* (*Young and Wise*). In 2003, she started to teach at the Amsterdamse Jeugd Theater School, training young performers in theatre to philosophise. Over the past decade Thecla Rondhuis gave numerous public presentations and interviews on TV and Radio, on the philosophical quality of children and youngsters' thinking patterns.