

Pays-Bas : measuring the philosophical quality of discussions in education

Thecla Rondhuis, PhD in philosophy, Scientific research at Utrecht University (formerly) and RIAGG Haarlem (thecla.rondhuis@planet.nl)

In philosophical education we aim to encourage and foster philosophically qualified thinking patterns. When dealing with philosophical questions, adolescents sometimes demonstrate remarkable authentic thinking patterns. But what makes such thinking patterns philosophically qualified? Can such quality be measured and scientifically assessed? And when are group discussions philosophically qualified? Intrigued by these questions, thinking patterns of 11 - 16 years olds expressed in philosophical discussions were investigated systematically. This article describes the scientific approach in addressing these questions and provides some results of the research. Following a theoretical search to characterise the 'mental activity' of philosophy, six philosophical indicators were uncovered by registering and measuring 14,393 concrete utterances of youngsters in 70 standardised philosophical discussions. Together, they construct indices of philosophical quality which proved to be objective, reliable and valid measures.

Main features of philosopher's activity

'Philosophical quality' strongly relates to concepts of 'philosophy', 'philosophising', and 'wisdom'. Attempts to define or characterise philosophy revealed a wide spectrum of results. However, philosophers are quite unanimous about the boundaries of their discipline.

Philosophical concepts and the activity of philosophising have evolved through history. 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. Exercises in thinking, based on energetic processes of questioning and answering, are still represented by dialogues called after Socrates. Here, in pursuing the truth, analytically and disinterestedly, philosophy arrived at its typical characteristic inclining towards indecisiveness and uncertainty.

A common denominator of philosopher's activity is its conception as the independent production of consistent thought, avoiding dogmas, certainties or definitive judgements. Concentrating on philosophy as a mental activity, contemporary philosophers identify similar characteristics of their discipline, from Husserl (1968), recommending suspending judgement and 'bracketing' our natural attitude towards the world, to Dewey (1966), claiming an epistemological and moral fallibilism and from Sellars (1963), describing philosophy as "the study of how things, in the broadest sense of the term, hang together, in the broadest sense of the term" to Merleau-Ponty (1967), emphasising our ambiguous mode of existence and our unwillingness for acquiescing in definite conclusions. So, common philosophical denominators may be found as well in joint historical threads and in

similarities between explicit views on philosophising of well-known 20th century philosophers, originating from widely diverse orientations (Rondhuis & van der Leeuw, 2000):

- Analysing and reasoning qualities
- Qualities detecting ambiguities, vagueness, uncertainty and borderline explorations
- Clarification of everyday experiences in moving smoothly from theory to practice and vice versa.

Empirical studies on wisdom (Sternberg, 2003; Brugman, 2000) revealed similar qualities, referring to individual human characteristics, with a strong involvement in life pragmatics, and a penchant for dynamic interaction between knowing and experiencing. However, philosophical and psychological approaches (wisdom) do not seem mutually compatible. As none of these can be ignored in the search for philosophical quality, the author believes that both should be placed under a joint umbrella, addressing thematic developments as well as empirical investigation of individual performances.

Thinking patterns of adolescents

Conceptual considerations derived from these main features of philosophy should be matched with observations of real life expressions uttered by adolescents while dealing with philosophical questions. Their thinking patterns cover a wide spectrum of cognitive qualities from sensitivity to ambiguities and reasoning qualities to explorative and tentative behaviour, trying to 'capture' the unknown, proceeding along their own non-predictable mental moves.

Examples of adolescents' thinking patterns

Sometimes, children's thinking patterns reflect classical philosophical ideas, bridging between real life experiences to domains of abstractions and vice versa. When a ten year-old girl was asked to clean the mess in her room, she dropped silent while considering this question she replied: 'Does a mess exist if nobody sees it?' Although the child was not aware of the great movements within epistemology, her question obviously but unintentionally points to the discussion between rationalism and empiricism. Whether something exists due to its observable qualities or just because of its logical inevitability to exist, is perceived as one of the most exiting questions by schoolchildren. Sometimes, they reason beyond presupposed contexts and expectations, accepting the independently growing conclusions of thinking trials readily, as may be observed in the comments of a boy (12) on the record of a drunk, who was drinking many bottles of wine (Saint-Exupéry's: *The Little Prince*, XII): 'If a drunk keeps on drinking, he will become old' (Rondhuis, 2001). Such thinking patterns may be right in a linguistic logical sense and at the same time wrong according to some practical experience. Some trains of thought construct circular or contradictory arguments or uncover hidden standards, as can be inferred from the suggestion of a twelve years old boy to make things bigger by placing them under a microscope and subsequently inventing a formula to assess the expanded version as the real (Rondhuis, 2001). Although this suggestion seems ridiculous, it may

turn out that only different standards were employed to qualify proportions of objects. Should size be measured by eye, by touch, or by abstractions? Perhaps Daniel Dennett (1993) experienced a similar fascination 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?'

Many issues challenge the brain powers of adolescents, ranging from stories to experiences that can be traced back to philosophical sources. The dynamics of thinking patterns in a discussion of 11 to 12 years olds dealing with a philosophical question born from contemplating the possibility of going further endlessly is described below.

Andrea: In a train you can keep on travelling, pass 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.

The text fragment demonstrates that the philosophical quality of the discussion exceeds reasoning quality, correct use of arguments, and presentation of valued knowledge or ideas. As Russell (1981) pointed out, philosophy is the no-man's-land between theology and sciences (i.e., all definitive knowledge). Philosophy is vulnerable to all attacks from the rational as well as from passion, astonishing, and imparting a degree of spiritual and intellectual uncertainty to the researcher. Many thinking patterns exposed by youngsters addressing philosophical questions share this intellectual turmoil.

Philosophical content

Although the discussion's starting question tries to identify the concept of endlessness, other themes are explored as well. The identity of a road is elaborated by emphasising its name (language), its observable and textural appearance, and rules attached to that name. The description of a road overlaps with its prescription, as embodied in a set of rules. The discussion offers an opportunity to discern differences between logical, conventional and practical consequences of using the concept road. Children and adolescents explore a variety of perspectives concerning endlessly passing tracks, searching for direction, movement and final goals. Simultaneously, the concept infinity is explored. Infinity addresses a theory of endlessness and an idea of everlastingness, stressing the meaning of going on, with its temporal and spatial components. Andrea explores practical experiences of endlessness and moves to circles and networks. Relativity is introduced. In Paco's next sentence a smooth transition from endlessness in space into endlessness in time is elaborated into immortality.

Children's answer-finding procedures

Through speculative utterances, the conventional meaning of endlessness with respect to road, seems to be stretched and to slip into figurative meanings. Such trial may lead into vagueness, but it also creates an opportunity for concept expansion. Vagueness strips concepts of their rigid meanings, constrained by correctness. Moving the meaning of an endless road into vagueness, challenges the thinking about space and time as through metonymic and metaphoric use of the concept. However, this does not imply that children are conscious of the wide meaning framework. Many sentences end with a question, some represent attempts at answering. These are expressions of tentative behaviour, often emphasised by the use of maybe. The number of suggestions and fantasising thoughts also exposes a high degree of creativity and attempting a solution for emerging problems. A new word is tried out: You just call it endlessly (originally, a new Dutch word was invented). Finally, sensitivity to ambiguity is expressed in: maybe yes but maybe not.

By joining in such trains of thought, performing philosophically qualified thinking patterns, children and adolescents are open and frank while running the risk of displaying attitudes vulnerable to criticism. They see no problem in disagreeing with themselves, are willing to avoid certainties or dogmas, and suspend generalised judgements. Vagueness and ambiguity, are the pre-eminent drivers in the process of evolving philosophically qualified concepts and ideas. In the empirical investigations into philosophical features of adolescents' discourse (Rondhuis, 2005), these are

treated as observable and countable occurrences. By evaluating adolescents' thinking patterns as a collection of such countable occurrences, potential discrepancies between theoretical and empirical psychological approaches are bridged.

Assessing philosophical quality

To uncover the philosophical quality of observable thinking patterns a conceptual framework and a measuring instrument are needed. The first is realised by developing six clearly detectable indicators, the second by a standardised format for philosophical discussions: the tetralogue. In a tetralogue, chaired by a qualified expert, four participants exchange their trains of thought ignited by a voluntarily chosen philosophical key question.

Philosophical indicators are determined by grouping of philosophical meaning components stored in linguistic expressions of youngsters. In its most simple mode, these comprise specific words, as verbs of modality or causal conjunctions. More complicated modes of expression are uses of metaphors, unconventional syntactical moves, ironic ways of expressing, and those as may be noticed from the discourse above. A variety of semantic, pragmatic, and syntactical features are collected to identify philosophically qualified thinking patterns (Rondhuis, 2005). They may differ in quality and are not detectable in one dimension only. However, all of them grasp philosophical meaning components. Following 'Grounded Theory' (Glaser & Straus, 1967), such components were grouped into indicator categories. Once indicator categories were saturated and no new phenomena of expressions appeared, six philosophical quality indicators were identified and assessed. Each indicator covers a group of linguistic expressions; together, they construct the conceptual framework for assessing philosophical qualities.

Six philosophical indicators:

- Indecisive thinking (Idt) denoting a mental inquietude or uneasiness, sensitivity to ambiguity and uncertainty, and awareness of the multiple ways to understand events.
- Openness (Op) denoting wonderment, the apparent readiness to meet the unknown, and denial to accept views without further ado.
- Tentative behaviour (Te) pointing to an attitude of mental experimenting and trying out.
- Epistemic position (Ep) denoting speaker's position with respect to the propositional content of his utterance. Sometimes it shows reflection, sometimes it depicts detachment between speaker and his declaration.
- Reasoning quality (Re) indicating all trials of analysing and reasoning, no matter whether argumentation is correct or false.
- Anecdotal quality (An) referring to the upholstering of ideas and concepts with real-life stories and experiences.

Some indicators like indecisive thinking and Openness demonstrate divergence in thinking patterns, while expressions of tentative behaviour converge. Anecdotal quality does not represent individual qualities, but rather depicts philosophical qualities in the context of the entire discussion only. All indicators can be attributed to oral expressions only if they are meant as such and not used as fillers.

Numbers, parameters, and statistical results

Occurrences of these indicators are registered, scored, and counted from youngsters' utterances in tetralogues. In *Philosophical Talent* (Rondhuis, 2005), 70 such tetralogues are recorded (video- and audio taped) with 215 participants in two age categories: 11 to 13 and 14 to 16 years old. Furthermore, two educational levels for participants were registered: high and low and another two categories concerning the regularity of life course to search for the influence of exceptional circumstances (physical disability, imprisonment, or living without natural parents) on the philosophical quality of thinking patterns. After formatting, the 70 tetralogues revealed a total of 14,393 utterances. Objectivity and reliability of the indicator assessment and of the scoring method was demonstrated by proving an inter-rater agreement (average Cohen's Kappa: 0.78), and by good split-half ($r = 0.841$) and test-retest ($r = 0.531$) coefficients. These results prove that random factors do not substantially influence the tetralogue's total indicator score. Moreover, the five assessed indicators judging individual performances appear to be internally consistent with a homogeneity alpha of 0.8.

Numerical indices to represent the philosophical quality of individuals and groups

Two numerical indices were constructed: pq and PQ, as separate indicator frequencies do not simply reflect philosophical qualities. The pq index expresses the philosophical quality of individually performed thinking patterns and is basically constructed of a balanced ratio of five indicator frequencies. The PQ index refers to group performances (tetralogues), measuring the philosophical power of collectively performed thinking patterns exceeding that of individual contributions. In the latter, anecdotal qualities and dialogue events of jointly generated, qualified combinations of indicator frequencies were taken into account as well. For example, the single presence of Ep, like in: "I think the car is green", is relatively worthless, while some combinations of indicators, such as Idt and Re signify a higher degree of philosophical quality than the sum of these indicator frequencies only.

Testing the validity of pq and PQ indices

To check the plausibility of the hypothesis that PQ and pq indices really pick up philosophically qualified thinking patterns, the indices are empirically related to independently obtained quality measures. At an early stage, all tetralogues were judged on a [0-5] point-scale by experts based on video registrations only while taking relevant overview criteria into consideration (Rondhuis, 2005). Calculated PQ indices are compared with these earlier obtained rough estimates. A Pearson correlation of $r = 0.541$ proves convergence between PQ indices and expert estimates. This correlation

is low enough to show distinctiveness between the two measures, for the PQ index is the result of objective, detailed, theory-based procedures and is expected to out-perform any crude judgment.

Indices of individual performances (pq) of 25 participants are compared with pq indices of the same participants in different tetralogues resulting in a correlation coefficient of $r = 0.611$. This correlation demonstrates considerable convergence of individual performances over two tetralogues and supports the hypothesis that pq indices do characterise individual performances. If this observed trans-situational convergence was 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. However, no significant correlation could be demonstrated between pq indices of participants in the first tetralogue with pq indices of a random selection of co-participants in the second tetralogue. This provides strong evidence for the pq index to reflect individual characteristics.

These results show that pq and PQ indices meet the prerequisites for being valid measures representing philosophically qualified thinking patterns, of an individual (pq) and a group (PQ) and that the main aim of this study was met by successfully corroborating the construct validity of the measurements.

Participant's characteristics in relation to their philosophical quality

The project also explores philosophical quality in relation to participant characteristics, serving two goals mainly: it is of psychological interest to put the individual philosophical quality on record in relation to well known cognitive and attitudinal qualities, but most of all, it is to corroborate the construct validity of the measurements. This was realised by checking whether theoretically expected relations from the nomological network can be replicated empirically. The nomological network includes a variety of individual and group characteristics. Among them individual characteristics that were measured in advance: educational level, word fluency, non verbal intelligence, five different personality traits, age, gender, and the experience of a regular or an irregular life course. All tetralogue participants were tested on personality traits by the Neo FFI (McCrea & Costa, 1990) and on non verbal intelligence by the Raven test (Raven, Court & Raven, 1977). They were asked for their Grade Point Average for language performance (GPA) and did complete a questionnaire with biographic details. The characteristics construct a network in which indices of philosophical quality are assumed to operate.

As expected on theoretical grounds, several links between philosophical quality and its performer could be identified. Indices of pq appear to be correlated significantly and positively with openness to experience as personality trait ($r = 0.271$) and with participant's educational level ($r = 0.441$). It is a matter of fact that the average philosophical quality in grammar schools is significantly higher than that in schools of a lower educational level. Indices of pq also correlate with intelligence (non verbal: $r = 0.152$, verbal: $r = 0.201$), but to a limited extent, indicating a difference between philosophical qualified and convergent thinking patterns. At the same time, no relation between philosophical quality and age is found, neither between philosophical quality and irregularity in life course.

Although, age variance was affected by the transition from primary to secondary school. The absence of presupposed relations with respect to an irregular life course (individual) or heterogeneity in life courses (in groups) may be attributed -at least partly, to the apparent impossibility of determining the characteristics of a 'regular' life course.

There are indications that individual philosophical quality may be related to habitual environmental circumstances that may hone, feed or discourage this quality (Rondhuis, 2005). 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 talent.

Philosophical quality "vis-à-vis" psychological criteria

Matthews (1994) noted the authenticity of children's thinking patterns recognised as philosophically qualified by professional philosophers. He compared these patterns with Piaget's stages of cognitive development. He discovered that authenticity and inventiveness do not fit presupposition of evaluative judgement to which the concept of childhood is submitted. Philosophically speaking, children's thinking patterns must be evaluated openly and without any goal orientation, presupposing a stage of maturity mastering cognitive operations (including moral and political correct concepts). Definitive answers and definite assessment of the right and wrong do not exist in philosophy. Do colours exist when nobody can see them? From a philosophical point of view, even physical truths are open to alternative interpretations. Children's thinking patterns may develop beyond well-defined steps, and sometimes they are even 'successful' by accident. Conventional developmental psychologists consider children's thinking patterns related to the cognitive human development into correct thinking according to norms assessed for adults. However, neither a development into correct concepts, nor a cognitive development into maturity is taken for granted in order to do justice to a level of cognitive inquietude and of being aware of the tentative and arbitrary nature of philosophically qualified thinking patterns.

Philosophical quality in philosophical education

Assessment of philosophical quality relates at least to three educational issues: 1) choice of themes; 2) discussion chair; 3) evaluation criteria.

Philosophical themes

Once the existence of a philosophical quality had been assessed, it may be questioned if that would depend on the philosophical theme discussed in response to a key question. Non-philosophical discussions, such as about social dilemmas relevant to the age of the participants, may show similar indices for pq and PQ as tetralogues as well. The key questions of all tetralogues were voluntarily selected by the four participants. Following the selected themes, all tetralogues were assigned to classic philosophical categories: metaphysics and epistemology, anthropology, ethics, and topics dealing with meaning and demarcation problems. Outcomes of individually (pq) or group (PQ)

performed philosophical qualities were checked against these categories. As is found in Philosophical Talent (Rondhuis, 2005), all categories of thematically grouped tetralogues show similar patterns of indicator frequencies. Different themes do not lead to significant differences in pq and PQ indices. Also, measurements of non-philosophical discussions in tetralogue format show lower indicator frequencies than are found in tetralogues on philosophical topics on average. The difference between philosophical and non-philosophical discussions appears to be not a difference of kind but of degree. Since tetralogue's philosophical qualifications (PQ) are unrelated to the four philosophical categories, the general nature of the conceptual framework and of the tetralogue as an instrument of philosophical enquiry is demonstrated, irrespective of its theme.

Nevertheless, differences can be noticed between participant preferences for key questions and themes: anthropological key questions appeared to be more popular among mixed groups with girls and boys than among males only, and more among the low educated than among higher educated youngsters. Female groups and higher educated participants preferred ethical themes.

Traditionally, some philosophical themes are exclusively connected with existential experiences, with thinking about life, death and God. Our hypothesis was that imprisoned, adopted and disabled youngsters would struggle with their delicate life situation, that they would ponder more than their regular peers, and would therefore demonstrate a substantially higher philosophical quality. This expectation was, however, not supported by the data, possibly because of poor assessment of the life courses of 'regular' youngsters. Nevertheless, some unique thinking patterns in philosophical discussions with 'irregulars' were registered. A severely physically disabled boy (Spastic Tetraparasis) for example, indicated that his handicap formed part of his identity. Therefore, he reasoned, his difficulty to communicate and contribute to society did not hamper him. Even if he was given a free choice, he declared that his handicapped existence, including his continuous dependence on a wheelchair and on many helpers and aides, was the best form of life imaginable. Also, philosophical discussions with imprisoned youngsters showed remarkable existential qualities. The predominantly Moslem boys kept in a closed judicial institution eagerly and openly discussed God's authority and power, but hated discussing issues as 'conscience' and inner 'voices'. Although touching outstanding existential questions, no high PQ's could be identified in the exemplified discussions. The ease of equating remarkable existential topics with a high philosophical quality may be credited to the discrepancy of presuppositions between participant and observer.

Rhetoric skills and debating classes

Although, rhetoric is part of philosophy, it chiefly focuses on reasoning skills. The main difference between tetralogues and debating sessions is that the latter aims at persuasion and the first at clarification. In case of debating, discussions develop more and more towards established opinions while addressing a decreasing number of questions. This performance contrasts with the development of qualified philosophical thinking patterns in tetralogues.

Chairing philosophical discussions

Tetralogues are chaired by qualified experts bound by predetermined interference rules. As the chair seems to play a role in both design and implementation of tetralogues, one might question the chair's influence on the PQ. This is explored by dividing chair persons according to their background (philosophical or psychological) and their experience (students or scholars). All received instructions and training. No significant differences between PQ indices in relation to these separated groups chair persons could be demonstrated. 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 that appears to depend on participants' individual traits.

Evaluation criteria

The relevance of the objective, reliable, and valid assessment of indices of pq and PQ is their discriminative power. Proving the existence of a distinctive philosophical quality, pq and PQ indices allow comparison between performances of philosophically qualified thinking patterns in philosophical discussions with adolescents. They may prove helpful in evaluating the philosophical quality of papers, essays and exam performances in philosophy. Traditionally, such evaluation focuses on convergent thinking patterns, qualified knowledge, language achievement, and supposed abilities to understand underlying principles. These valuable qualities cannot be measured by the indices presented here. However, application of pq and PQ measurements can contribute to a more complete evaluation of typical philosophical thinking patterns because they reveal specific qualities such as sensitivity to ambiguities. Adaptation of indices for philosophical quality might be realised by admitting discussion qualities or recognising the alternation of contrasting ideas to individual performances. Likewise, alternation of anecdotes and trials to exceed these through regressive abstraction should be identified in individually written papers and essays.

What is the relevance of philosophical quality?

As assessed in this study, philosophical quality refers to a complex of cognitive factors, including sensitivity to ambiguities and uncertainty, reasoning and analysing qualities, and the ability to shift smoothly from concrete to abstract and vice versa. This complex differs from cognitive qualities of convergent thinking that can be measured by conventional tests. The assessment of a philosophical quality determines philosopher's activity as a separate domain of cognitive behaviour. The outcomes can be extended for evaluation and selection purposes.

Paradox

Measuring and assessing supposedly intangible philosophical qualities might be perceived as paradoxical, as philosophy is supposed to be devoid of any clear, final goal. But we cannot describe philosophy, evaluate philosophical essays or discussions, or remove philosophy from interminable

thinking patterns without some kind of measurements. Therefore, philosophical thinking patterns would need to be described in terms that allow comparison, including a measurement tool. Notwithstanding the empirically achieved results reflecting valid approximations to the philosophical quality concept, they do not pretend to cover the entire concept.

Conclusion

Our studies on philosophical talent demonstrate that philosophical quality of thinking patterns does exist. Moreover, it identifies an exclusive cognitive endowment, different from convergent thinking patterns and can be measured in a scientifically valid way for groups (PQ) and for individuals (pq). Philosophical quality proves to be a stable, personal quality driving authentic thinking by adolescents when dealing with philosophical topics. Finally, it provides a mechanism to quantify earlier qualitative approaches of philosophically qualified thinking patterns of adolescents.

(1) means a significance level of $p < 0.01$

(2) means a significance level of $p < 0.05$