DOES THE CONDUCT OF CRITICAL THINKING ADD TO THE IMPACT HIGHER EDUCATED YOUNG PROFESSIONALS MAKE?

by

ing. Boudewijn de Graaf MSc CIPM CIPP/e

DISSERTATION
Presented to the Swiss School of Business and Management Geneva
In Partial Fulfillment
Of the Requirements
For the Degree

DOCTOR OF BUSINESS ADMINISTRATION

SWISS SCHOOL OF BUSINESS AND MANAGEMENT GENEVA

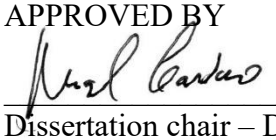
<JUNE, 2025>

DOES THE CONDUCT OF CRITICAL THINKING ADD TO THE IMPACT HIGHER EDUCATED YOUNG PROFESSIONALS MAKE?

by

ing. Boudewijn de Graaf MSc CIPM CIPP/e

APPROVED BY



Dissertation chair – Dr. Gualdino Cardoso

RECEIVED/APPROVED BY:

Admissions Director

Dedication

Conducting Critical Thinking is not always as fun as a party. It's more like the bright fluorescent light after a party....it's enlightening, but you don't always see what you would have liked to see and sometimes you see what you wouldn't have wanted to see.

This work is dedicated to all people who have the bravery to find out what really was, is and can be “in here” and “out there” and use these insights to make positive Impact on life itself.

.

Acknowledgements

It's impossible for me to call all the people who helped me during my DBA by name here. Not only because the amount of people would result in an unworkable movie-credit like list. Also, I would probably miss a name. And with a list that long, missing someone would possibly be painful for this supporter; and it would be surely utterly embarrassing for me. So, "thanks folks"!

However, I would like to name a few people that stood out. First of all, I would like to thank my promotor, Professor Velimir Srića PhD MBA MS for the trust and support he gave me during this journey. I thank Mr. Abdel Yalte for giving me the opportunity to let this long-held wish come true. I would also like to thank a diversity of colleagues and students at the Rotterdam Business University of Applied Sciences -and some others outside this institute- for the support and taking part of the pre-tests of measurement tools. Also, I want to thank the Young Professionals and their managers: thanks for your input!

My special thanks go out to Daisy for close reading the earlier concepts and to Liselotte who did a great job in finding Young Professionals and their managers when I needed it most.

Elspeet, March 2025

Boudewijn de Graaf

ABSTRACT
DOES THE CONDUCT OF CRITICAL THINKING ADD TO THE IMPACT HIGHER EDUCATED YOUNG PROFESSIONALS MAKE?

ing. Boudewijn de Graaf Msc CIPM CIPP/e
2025

Dissertation Chair: Dr. Gualdino Cardoso
Co-Chair: Dr. Ljiljana Kukec

This dissertation investigates whether the conduct of Critical Thinking (CT) enhances the Impact that higher-educated Young Professionals (YPs) make at the start of their careers. While CT is widely promoted in the curricula of higher education institutes (HEIs), little research was available if CT, conducted by YPs really results in making Impact. This study introduces a new conceptual model integrating CT, Mental Models, and Impact. A perception-based survey including a custom CT ability test among 29 Dutch HEI graduates and 11 of their managers measured the CT-ability, behavioral conduct and Impact of YPs. Results reveal a statistically significant, strong correlation between Friendly Voiced outward-directed CT conduct—not merely CT ability—and positive Impact. Findings suggest that developing CT conduct early during higher education and incorporating the development of Friendly Voiced outward-directed CT through application in later years of higher education through real-life assignments can substantially enhance a YP's effectiveness. Additionally, discrepancies were found between YPs' perceived freedom to exercise CT and managers' assumptions, indicating that organizations could foster greater Impact by mentoring and encouraging Friendly Voiced outward-directed CT behavior. The research offers practical recommendations for higher education institutions and employers and calls for further exploration into longitudinal effects and broader organizational contexts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I:
INTRODUCTION

1.1 Preamble

The world is changing rapidly. To make "Impact" and keep up with the exponential rate of innovations, Critical Thinking skills are broadly considered as a competence fresh graduate of Higer Education Institutes (HEIs) should master (World Economic Forum, 2016). However, applying Critical Thinking to this claim itself reveals that scientific evidence that conducting Critical Thinking leads to making Impact is slim, and the logic behind it is not as self-explanatory as it seems. Does Critical Thinking, conducted by Young Professionals recently graduated from Dutch Higher Education Institutes, indeed lead to Impact?

1.2 Research Problem

One of the main goals of Dutch HEIs[1] is to help students to develop themselves to YPs that can make Impact in our society. To accomplish this, they are not only introduced to the particular field of expertise of their chosen study. They are also introduced to "General competencies" they should master, and CT is one of them (Tadema, 2020). However, there is little research available on if conducting CT does lead to Impact, or if there are some modifiable conditions that can enhance the positive effect of conducting CT by YPs. A lack of knowledge makes considering alternatives and choosing between them more difficult.

[1] In this study Higher Education Institutes will be abbreviated as HEIs, Critical Thinking as CT and Young Professionals as YP's

This knowledge gap challenges HEIs, as they aim for their students to make Impact in society and contribute to the Social Development Goals (EFDM, 2023). HEIs can possibly enhance their curricula.

This knowledge gap is also a problem to the YPs, as they might be better equipped to make Impact at the start of their career.

This knowledge gap is also a problem to the employers of the YPs, as they should want employees that (help them to) make Impact.

And last but not least, this knowledge gap is a problem to society as a whole because it pays for, and wants to maximize the benefit from, investment in higher education.

1.3 Purpose of the Research

The objective of this study was:

1) Find evidence to falsify or support the claim that CT conducted by YPs at the start of their careers correlates positive with the Impact they make. And, if this was the case
2) Formulate actions for HEIs and Employers regarding learning and conducting CT during and after the study of YPs that will most probably add to the Impact they make at the start of their career

1.4 Significance of the Study

The results of this study are valuable for the people responsible for the curricula of the HEIs, especially the Rotterdam Business School -RBS-, part of the Rotterdam University of Applied Sciences, where most of the research will take place. It might lead to shifts in what parts of the CT education may need more emphasis, what should be

added and what could be omitted. Also, it is valuable for organizations where YPs start their career, as it might lead to programs to maximize their Impact. And this study also adds to the closure of the knowledge gap of the relationship between CT and Impact.

1.5 Research Questions

The main research question was: “Is there a relationship between Critical Thinking by HEI graduated Young Professionals and to the Impact they make at the start of their careers?”. And, if so, the second research question became relevant “Does Critical Thinking by HEI graduated Young Professionals lead to making Impact at the start of their careers? “. The third and final question was “What actions can HEIs and Employers take regarding learning and conducting CT that will most probably lead to the Impact YPs make at the start of their career?”

1.6 Reading Guide

The introduction in this chapter is followed by a Literature Review (Chapter II), where the existing knowledge is discussed. Based on this knowledge, Chapter III describes the followed methodology. In this chapter various topics are covered, amongst them the operationalization of the theoretical constructs, the breakdown of the research questions, the developed instrumentation and the participant selection. Chapter IV covers the obtained results. In Chapter V the interpretation of this results is given. In Chapter VI the research questions are answered, and recommendations are given. Used terms and definitions are other appendices can be found at the end of this report.

The next chapter gives the theoretical background that was used as the starting point of this study.

CHAPTER II:
LITERATURE REVIEW

In this chapter the theoretical background of this study is discussed. In paragraph 2.1 insights regarding Critical Thinking (CT) are given. Paragraph 2.2 is about Impact. Paragraph 2.3 summarizes the given theory and places it in the context of this research.

2.1 Critical Thinking

In the upcoming sub-paragraphs views on CT are discussed, the relation with Mental Models is made, common pitfalls are discussed as well as the way CT can be measured.

2.1.1 Definitions of, and views on, Critical Thinking

There are many different definitions of what CT "is". Lai (2011) argues, based on an extensive literature study, three main approaches exist: a philosophical approach, a psychological approach and an educational approach. The philosophical approach puts emphasis on how an "ideal" person should think, poses standards of "good thought" critical thinkers should have, using the formal ways of logic. Critics point out that this might be unrealistic in real world problems. The psychological approach with its roots in behavioral research puts emphasis on how people actually think and reason, versus the way they do it under ideal conditions. Critics point out that this approach simplifies CT and narrows it down to simple steps and instructions, just because this way the thinking behavior is measurable. The educational approach is merely interested in teaching and assessing CT as a mean to an end. Critics argue that this approach is quite vague about what CT is precisely. The categorization of the origin of Critical Thinking is supported by Lewis & Smith (1993).

Benjamin Bloom's taxonomy, particularly the levels of analysis, synthesis, and evaluation, is often used to represent Critical Thinking (CT) in the educational field (Nentl and Zietlow, 2008). Some researchers in the educational field also consider CT as a way for making students well-rounded persons, known as “Bildung” (Beck, 2019). CT can also be considered as a mean of fostering active citizenship and sustainable development (Uribe-Enciso, Uribe-Enciso and Vargas-Daza, 2017), and as a mean to harden society against malformation (Rotilă, 2022)

The three above mentioned origins of CT definitions have common grounds in defining CT as analyzing, making inferences, evaluating/judging and making decisions/solving problems. “Good” CT involves having the ability to conduct clear reasoning, the willingness to take dispositions and the virtue to live up to ethical standards (Lai, 2011).

Phan (2011) states, based on other scholars work, that CT is reflective thinking of an higher order, with Habitual action, Understanding and Reflection in ascending but lower orders of complexity. “Habitual action is a mechanical and automatic activity that is performed with little conscious thought. Understanding is learning and reading without relating to other situations. Reflection concerns active, persistent, and careful consideration of any assumptions or beliefs grounded in consciousness. Critical thinking is considered a higher level of reflective thinking that involves individuals becoming more aware of why they perceive things, the way they feel, the way they act, and what they do” (Phan, 2011, p. 285). The study of D’Alessio et al (2019) suggests that CT has a positive Impact on the academic performance of MBA students. The adoption of deep-learning strategies is positively correlated with the slope of adapting CT approaches. In a longitudinal 2-year study amongst 319 students at a local university Phan (2011) found,

to his surprise, evidence that an increase in deep-learning "understanding" strategies (as opposed to surface-learning "reproducing" strategies) is associated with a decrease in conducting CT. Phan suggests that his findings regarding the decrease of CT through the years might be influenced by inconsistencies in the "constructive alignment" of the curriculum, stating previous research findings showed a strong bi-directional association between CT and deep learning. Another explanation of the finding of Phan may be found in the stages in the Conscious/Unconscious and Competence/Incompetence matrix (Cannon, Feinstein and Friesen, 2010, p. 177). As students become more competent in their field of expertise, the need for CT may diminish as their actions no longer depend on conscious reasoning.

Critical thinkers are eager to understand the world as it is. A critical thinker "is willing to examine his/her beliefs, assumptions, and opinions and weigh them against facts." (Natale and Ricci, 2006, p. 3). Other authors describe the critical thinker in a similar way (Thonney and Montgomery, 2019). What is striking is the way most mentioned authors do not elaborate on how CT relates to "thinking" in general, what kinds of "thinking" can be recognized and what the nature of thinking is. As Peters states "The contemporary tendency reinforced by first generation cognitive psychology was to treat thinking ahistorically and aculturally as though physiology, brain structure and human evolution are all there is to say about thinking that is worthwhile or educationally significant" (Peters, 2007, p. 350). Peters recognizes three waves of philosophical revolutions and concludes that thinking is rather a social construction, rather than an individualist cognitivist construction. As we will see later in this report (paragraph 2.3) the researcher incorporates this "social dimension" and the narrowed cognitive interpretation in a construct of different dimensions and "layers" of Mental Models.

2.1.2 Critical Thinking and Mental Models

As adapting is one of the main elements of survival and prosperity both as an individual and as a group (Gould and Lloyd, 1999), the capacity to reason and use imagination may be key in the current success of human kind. “One of our biggest distinctions as a species, however, is our unique capacity to make counter-evolutionary choices (..). This includes the ability to adapt our mental models or mindset and in consequence of that also adapt our strategies and behaviors when, e.g., negotiating change, making decisions, or dealing with differences.” (Jordaan, 2019, p. 63). A consistent and coherent MM maximizes our evolutionary success. Realizing one’s MM is not internally consistent and coherent, causes existential stress to “the Self” that the brain wants to relieve. This triggers well-researched phenomena like cognitive dissonance and self-deception (Cline, 2016)

A MM can be defined as “(..) a “selective abstraction” of reality that you create and then carry around in your head. As big as some of our heads get, we still can’t fit reality in there. Instead, we have models of various aspects of reality. We simulate these models in order to “make meaning” out of what we’re experiencing, and also to help us arrive at decisions that inform our actions.” (Richmond, 2010, p. 2). MMs are therefore limited. “Mental models tend to be functional rather than complete or accurate representations of reality. A mental model is a simplified representation of reality that allows people to interact with the world. Because of cognitive limitations, it is neither possible nor desirable to represent every detail that may be found in reality. Aspects that are represented are influenced by a person’s goals and motives for constructing the mental model as well as their background knowledge or existing knowledge structures, (..). Mental models thus play a role in filtering incoming information.” (Jones *et al.*,

2011, p. 5). Incorrect MMs lead to insufficient problem solving (Thompson and Van Boven, 2003).

Thinking, reasoning and building up MMs are intertwined "I will define thinking as consisting of two activities: constructing mental models and then simulating them in order to draw conclusions and make decisions." (Richmond, 2010, p. 2). A MM makes reasoning possible: "A mental model is constructed in working memory and can then be run like a computer simulation allowing an individual to explore and test different possibilities mentally before acting. Working memory is the system responsible for selecting and manipulating information for the purpose of reasoning and learning. Changes made to a mental model in the simulation process represent what would happen if such changes took place in reality." (Jones *et al.*, 2011, p. 4). "Although the mental model concept has its share of critics (..) there is nevertheless convergent evidence that mental models are associated with differential success in reasoning and problem solving across varied domains." (Thompson and Van Boven, 2003, p. 388)

One way to validate and improve MMs is CT. "The essence of our theory is that critical thinking skill is exemplified by asking questions about alternative possibilities in order to achieve some objective. Asking and answering questions is a skill of dialogue. Alternative possibilities are represented by mental models. A process of questioning mental models is adopted because of its reliability for achieving the purposes of the participants within the available time." (Cohen, 2000, p. 3). However, CT is not the only way to build a valid MM. The current emphasis of ratio in techniques for building valid MMs is clear, but also criticized "despite this diversity, all the new methodologies and methods are still founded on principles of rational analysis, and approaches that go beyond this are marginalized. For instance, arts-based and theater methods are rarely mentioned in the literature on systems thinking, yet they can help people discover how

their value and boundary assumptions have roots in unconscious impulses and memories." (Midgley and Rajagopalan, 2020). The classic concept of a MM as a cognitive, rational set of entities and relations between them (Jones *et al.*, 2011) can be stretched. LeDoux formulates, based on research, an integrated model where stored "facts and concepts" lead to a "perceptual mental model", where at the same time stored "emotions, facts and concepts" and stored "Self facts and concepts" lead to an "emotional mental model". Thus, MMs can be perceptual and emotional. Perception, emotion and an active "Self" scheme all take part at translating what is sensed in conscious experiencing "perceiving" and "feeling" (LeDoux, 2020). His idea that the CT, perception, emotion and the "Self" are strongly intertwined is, in a way, a next step in line with an older concept. CT plays a central role in Cognitive Therapy (Knapp and Beck, 2008), using reasoning to correct ineffective thought patterns ("Cognitive distortions") and thus affecting emotional well-being. (Irwin and Bassham, 2003).

MMs can be defined at individual, team, organizational and society level: "Cultural models are discussed in a similar light to collective mental models and shared mental models in that they all refer to a degree of shared understanding among a group people." (Jones *et al.*, 2011, p. 3). "(..) these mental models seem to guide social behavior and perceptions of social situations." (Thompson and Van Boven, 2003, p. 389). Shared MMs tend to maximize the performance of groups (Thompson and Van Boven, 2003). On top of that, in an organizational culture where deviant ideas that challenge the common opinion are valued, CT improves the quality of the group decision, which in turn tends to improve organizational performance (Natale and Ricci, 2006)

CT improves the quality of decision making (Haase, 2010; Helsdingen *et al.*, 2010). CT may contribute to building valid MMs. "Thinking allows humans to make sense of, interpret, represent or model the world they experience, and to make predictions

about that world. It is therefore helpful to an organism with needs, objectives, and desires as it makes plans or otherwise attempts to accomplish those goals." (Balasubramanian and Fathima, 2011, p. 2). However, limited research of the nature and magnitude of the contribution of conducting CT to building a valid MM can be found. There is some research about how MMs can stimulate CT (Krejci, 1997), which is "the other way around". And, as stated before, there is extensive research in the field of Cognitive Therapy, suggesting that conducting Crtitical Thinking does enhance MMs (Beck, 2005). In Cognitive Therapy, people are made aware of distortions in their MMs. This way people can "fix" their MM, and thus come to "better" (emotional) reactions (Knapp and Beck, 2008). A language-pattern approach to challenge and enhance MMs is the "meta-model", which also can be used in qualitative research (Knight, 2012)

2.1.3 Critical Thinking: pitfalls and problems

In order to structure the Pitfalls and Problems, this paragraph is structured using the OMA=P formula. This formula states that Performance can only be expected when a person is having the Ability, Motivation and Opportunity to do so (Kellner, Cafferkey and Townsend, 2019). Pitfalls and problems for "good CT performance" can arise on all three aspects.

The ability to think critically may suffer from the lack of formal reasoning competence. Extensive lists of thinking errors are available in literature (Kahneman, 2013; Nikolopoulou, 2023). Systematic approaches to "challenge" thinking errors in MMs do exist (Knight, 2012; Jin *et al.*, 2022). The ability to think critically is also affected by inherent weaknesses from the MM(s) in use, as sensed stimuli may be false, incomplete, misinterpreted, ignored, "enriched" with assumptions based on prior knowledge, processed in oversimplified or biased models of reality and results may be in

conflict with the established model, triggering reactions as disbelief or loss-aversion (Thompson and Van Boven, 2003; Jones *et al.*, 2011; Jordaan, 2019; Dharani and April, 2022). For adolescents an extra hurdle occurs; until the age of 25 the human brain is still "work in progress". During adolescence the limbic system, which is related to survival, is more active than the Prefrontal cortex, which offers the capacity to conduct good judgement and correct behavior in social situations (Arain *et al.*, 2013).

The motivation to think critically may also be problematic. The argumentive theory states that our reasoning capabilities are merely a way to convince ourselves and others that our viewpoint is correct, and not in the first place to change our viewpoint based on facts (Mercier and Sperber, 2011). Evidence exits that reasoning capabilities are not used to their full potential in cases where facts are not in favor of the current MM, supporting the Identity-protective Cognition Thesis (Kahan *et al.*, 2013). CT is hard work. Nobel prize winner Kahneman argued that the ways of thinking of humans can be (artificially) divided in 2 systems (Kahneman, 2013). System 1, guided by impressions and feelings, is fast, automatic and effortless, leaning on heuristics and stereotypes, with as main function maintaining and updating the model of the personal world and keep a comprehensible, causal and coherent story. Even if it does not match the uncertainty in, and complexity of the "real" world. System 2 is hard, slow, and analytical. The division of thinking in 2 systems is highly efficient in terms of distributing scarce attention, however, the "automated" system 1 thinking is also the cause of a lot of thinking errors (Stojanović, 2013).

The opportunity to think critically may be problematic, as the social context may influence the conduct of CT. Burris (2012) found that employees who were challenging the status quo of generally accepted sets of practices, policies and strategic directions ("challenging voice") encounter more resistance then employees that proactively vent

ideas that support this status quo ("supporting voice"). Dungan et al (2015) argue that whistleblowing -an ultimate form of challenging the status quo- is a tradeoff between loyalty (which can be seen as a form of belonging to and supporting the organizational Self) and fairness (which can be considered as a form of criticizing the espoused norms within the organizational Self by comparing them with external standards). Personal traits that are positively correlated with the willingness to being non-conferment to the status quo of the organization, are a feeling of responsibility of your own actions (internal locus), a proactive personality and a firm position in the organization (greater occupational power as a product of increased pay, increased education and being male) (Dungan, Waytz and Young, 2015). Organizational factors that support the non-conformism are organizational support and encouragement for this behavior, knowledge of the proper avenues for reporting, and the amount of protection from retaliation. Another factor of the willingness to show non-conformism is when unethical behavior comes abruptly rather than slowly over time ("boiling frog" apologue). Finally, it is stated that variability in the degree of collectivism of the organizational culture affects the willingness to be deviant; individuals from Asian cultures are less willing to be a whistleblower than individuals from America. In order to enhance group decision making, increasing innovation, critical consideration of a group's ideas and adaptively reflect upon their overall objectives, organizations might focus on building community's that values constructive dissent while maintaining group loyalty (Dungan, Waytz and Young, 2015).

2.1.4 Measuring the Conduct of Critical Thinking

There are multiple methods to measure CT as a competence. Well-known tests are the "California CT Skills Test (Facione, 1990), the Cornell CT Tests (Ennis & Millman,

2005), the Ennis-Weir Critical Thinking Essay Test (Ennis & Weir, 1985), and the Watson-Glaser Critical Thinking Appraisal (Watson & Glaser, 1980)" (Lai, 2011, p. 38). The Halpern Critical Thinking Assessment is translated in Dutch (De Bie, Wilhelm and Van Der Meij, 2015). In their work forementioned De Bie et al. discuss a wide set of tests and argue that measurement should test for both the ability to think critically, and the willingness to do so.

Using the OMA model, it can be concluded that the above-mentioned tests all measure the ability to think critically, some also measure the motivation to think critically, but none of them measure the opportunity to think critically. Also, none of the mentioned tests measure the actual conduct of CT in a business context. At least some of the above-mentioned tests do not address recognizing thinking biases.

2.2 Impact

In the upcoming sub-paragraphs views on Impact are discussed as well as the way CT can be measured.

2.2.1 Impact: Definitions and Frameworks

The Cambridge dictionary defines Impact as "the strong effect or influence that something has on a situation or person", as well as "a powerful effect that something, especially something new, has on someone or something" (Cambridge Dictionary, 2023). "Impact is the extent to which the intervention has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects. Impact addresses the intervention's ultimate significance and potentially transformative effects – holistic and enduring changes in systems or norms. The Impact criterion goes beyond effectiveness and encourages consideration of the big "so what?" question. This is where

evaluators look at whether or not the intervention created change that really matters to people." (OECD, 2021, p. 11).

Impact can be expressed via indicators. "Indicators are signposts of change along the path to development. Indicators are what we observe in order to verify whether – or to what extent – it is true that progress is being made towards our goals, which define what we want to achieve." (Sandhu-Rojon, p. 4). She distinguishes three types of indicators: situational, outcome and output indicators: Situational Impact indicators provide a broad picture of the situation, outcome indicators are more detailed versions of situational indicators and output indicators assess progress against specific operational activities. Indicators may combine qualitative and quantitative observations and may be long-term or short-term oriented. Indicators can be built up by other indicators, thus creating an "index" for complex phenomena. However, "The word indicator has a scientific aura, but there is no overall accepted precise definition of this term that is free from a halo of associations." (Heink and Kowarik, 2010, p. 591). Gudmundsson et al (2010) share the same conclusion, however state that "An indicator is generally understood to be a tool or a method which can be used to mirror or measure something in a way that adequately represents what is being measured" (Gudmundsson, Joumard and Aschemann, 2010, p. 24). Although "An indicator is never assumed to provide a complete description of something." (Gudmundsson, Joumard and Aschemann, 2010, p. 24). Quality aspects of (sets of) indicators are: the set should be coherent and balanced and indicators themselves should be Relevant, Accurate, Important, Useful, Feasible, Credible, Valid and Distinct ('Selection of Indicators', no date). Impact indicators are useful in a framework where there is some "Theory of Change" ('How Does Theory of Change Work?', no date). Witnessing Impact is an indication that the actions, based on the Theory, are successful, and therefore the Theory of Change holds. A Theory of Change can be regarded as a MM

of how the reality “works”. Creating a Theory of Change is a step explicitly mentioned in the web-based Impact Measurement tool developed in commission of the Dutch Ministries of Social Affairs and Employment, Economic Affairs and Climate Policy and Foreign Affairs (n.b., 2020)

The keyword “Impact” scores 7.930.000 results on Google Scholar. In line with this finding, depending on the aim of an Actor and the field of interest, there is a broad range of potential frameworks that could be considered relevant for measuring Impact at personal, organizational and societal level (Arieli, Sagiv and Roccas, 2020; Errida and Lotfi, 2021; OECD, 2021; Devassy *et al.*, 2023; EFDM, 2023).

Even studies aiming at creating a taxonomy for indicators, tend to define a ”model” of a particular field of expertise (Domínguez *et al.*, 2019; van de Ven *et al.*, 2023).

Measurement of (social) Impact depends on what the Actor recognizes as important: “This means the exercise is in fact far from ‘neutral’, as the selection of impacts to be scrutinized depends to some extent on the prior assumptions of the assessors. The assessment exercise can be further complicated when there are conflicting views with regard to the actual objectives of social policies.” (Freudenburg, 2019, p. Sec1:39).

The focus of our research will be on the Impact young HEI-professionals can make in SME’s, using CT. SME’s make up for more than 99% of the companies in Europe but work with SDG’s in less proactive and ambitious ways. Their primary focus is related to day-to-day activities (Nygaard, Kokholm and Huulgaard, 2022). The attitude of the Hogeschool Rotterdam towards SDG’s however is quite different, as they want to make Impact on SDG related goals (Hogeschool Rotterdam, 2015). This leaves us with the notion that “measuring Impact” with an overarching single-truth set of fixed

indicators is quite impossible, as the definition of what "making Impact" is about, varies widely.

2.2.2 Measuring Impact: stepping backwards

Stepping backwards and considering the above-mentioned Impact-measurement systems at a more conceptual level, the measurement of Impact follows the same basic structure. a) the Actor poses/chooses/defines a model of the world, build on what phenomena in the real world are deemed relevant to him, and therefore defines (sometimes implicit) what the borders of the system are; b) the Actor chooses/defines indicators that "measure" the status of the relevant phenomena and probably chooses "ideal values/norms" for these indicators; c) the Actor chooses/defines and implements a measuring and reporting system.

Going through the afore mentioned systems, Impact-indicators may (not limitative) express: Timeframes (e.g. duration of the effect; or time between the starting moment of the action and the start of observing the effect); Effect sizes (ratio's, e.g. number of elements in the defined system that are affected against total number of elements in the system); Direction (e.g. moving towards or away from a benchmark), Relevance/weight (e.g. how "important" is the primary value of the Actor influenced by the phenomena that should it be optimized/minimized, or does it only influence peripheral phenomena and should it for example simply not pass a threshold).

Actors can be defined at individual, group, organizational and societal level. The definition of what "Impact" should encompass is defined by the idea of the "Self", the values and worldview of the Actor. Therefore, "Impact" is always relative to the MM of the Actor. In order to measure Impact across a wide variety of Actors, instead of using an overarching single-truth set of fixed indicators, it could be more fruitful to measure

relative changes in the top-x phenomena every Actor based on its MM defines as "important", on dimensions like Time, Effect-size, Direction and Relevance. This way the measurement of Impact can be to some extend standardized and normalized, even when the underlying MMs -and with that the definitions of Impact- differ.

2.3 Summary

CT has, to the best of the authors knowledge, not been studied thoroughly in relation to real world Impact. CT involves an Actor that conducts clear reasoning, the willingness to take dispositions and the virtue to live up to ethical standards, in order to see itself and the world "as it is". CT may be problematic to conduct, as the ability, motivation and opportunity may not be in favor of conducting it. Oversimplification, the tension between fast and slow thinking, compartimalization in subsystems, the reluctance of letting a so-far-effective MM go, filtering and goal-oriented "argumentive" reasoning may occur. Even the ability to reason may be negatively affected when evidence is not in favor of the "Self" image in the MM. Also, the physical immaturity of the brain, the personality of the young professional and the social context might influence the conduct of CT.

Based on (LeDoux, 2020) a simplified model of a MM of an Actor with the dimensions Feeling/Perceiving, Conscious/Unconscious and an image of the "Self" (see Figure 1) is made. Research on Cognitive Therapy suggests (amongst others) that CT can, via the quadrant Conscious/Perceiving, contribute to a consistent and coherent MM for the "Self" of the Actor. The concept of MMs of Actors can be applied from personal to societal level.

Figure 1:
Simplified Mental Model

There is no common agreed upon framework for measuring Impact. The measurement of real-world Impact is strongly connected to the goals of the Actor and his view on which phenomena in the real world are important and which are not. Stepping back from several detailed models of measuring Impact, an intermediate element between CT and (measurement of) Impact might be the concept of MMs. Focusing on the extend an effect occurs relative to the MM of the Actor, can "normalize" measurements in completely different situations and makes results comparable.

It might be argued that conducting CT may lead to "effective" MMs on personal, group, organizational and societal level, which in turn may steer actions that maximize Impact. This logic is shown graphically in a thinking model in Figure 2. The CT Opportunity, Motivation and Ability should lead to CT Conduct. Conducting CT affects the Conscious perceiving quadrant, which should lead to an enhanced Mental Model as a whole. This could apply to the Mental Models on Personal, Group/organization level, and even to the context level. This "better" Mental Model should lead to effects on the things deemed important by the Actors and is experienced as "Impact". There might be feedback loops over time, as Impact may also alter the Mental Models, and the conduct of CT. Feedback loops can be positive ("You did well, now please do more of the

same…") but also negative ("This was a disastrous adventure that we will never do again!").

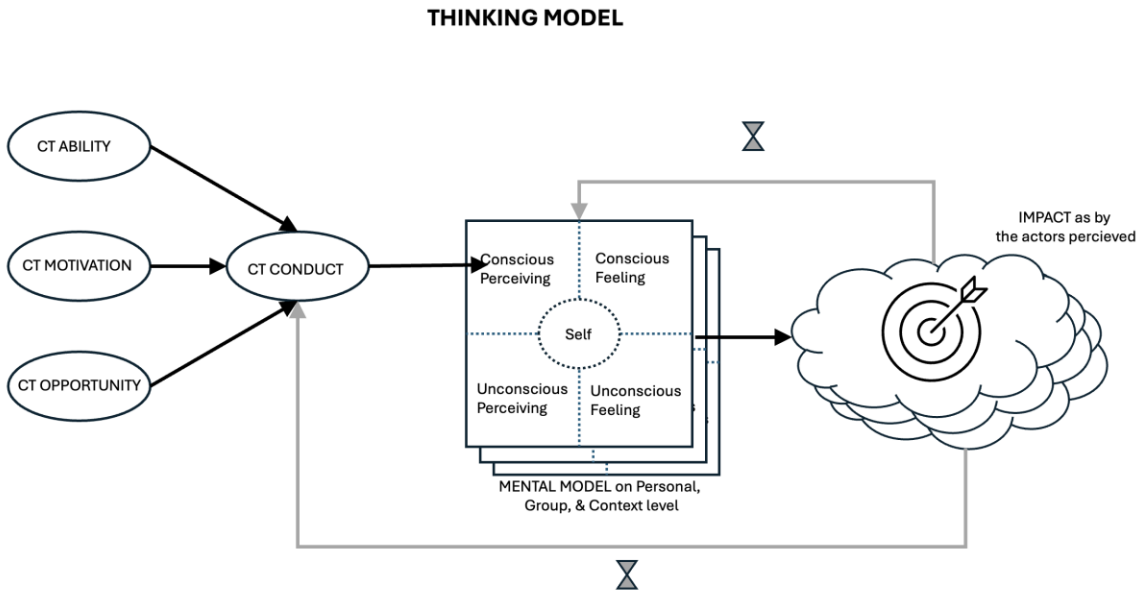


Figure 2:
Thinking Model

This research contributes to the "gap" in literature between CT and MMs, and at the same time may shed a new light on measuring Impact. Given the before described insights literature gives, the next chapter will be dedicated to how the research question was addressed.

CHAPTER III:
METHODOLOGY

3.1 Making the Research Problem Researchable

CT is supposed to be of great importance to make Impact. However, research on the Impact of CT is slim. Does the conduct of CT indeed lead to making Impact? Is there a correlation between CT and making Impact? Given the timeframe available and the working environment of the researcher (the Rotterdam Business school, Rotterdam University of Applied Sciences), the broad question if CT greatly adds to making Impact is narrowed down to finding out whether CT conducted by Young HEI graduated Professionals really has a positive Impact on the Impact they make in their working environment. But even this narrowed down question isn't straightforward to answer. As argued in paragraph 2.2 for example, "the" definition of Impact (and therefore the measurement of Impact) varies widely. As the RBS YPs most likely end up in completely different fields, none of the "standard" Impact Measurement Systems is usable as a one-size-fits-all measurement system.

The overarching questions mentioned in paragraph 1.5 leads to practical research questions:

- What are, for this study, the definitions of CT, Mental Models and Impact,
- How can these constructs be operationalized and
- What hypotheses will be tested using these constructs?

The answers on the above questions give input to two main research questions “Is there a relationship between Critical Thinking by HEI graduated Young Professionals and to the Impact they make at the start of their careers?”. And, if so, “Does Critical Thinking by HEI graduated Young Professionals lead to making Impact at the start of their careers?”.

This chapter addresses, based on the literature review in Chapter II, several topics. First working definitions of CT, Impact and Mental Models are given. Then the logic behind the supposed relationships of those concepts is addressed. Als the measurement of these concepts and relationships are described.

Chapter IV describes the outcomes of the measurements as described in this Chapter III. Chapter V gives an interpretation of the measurements and Chapter VI answers the third and final question “What actions can HEIs and Employers take regarding learning and conducting CT that will most probably lead to the Impact YPs make at the start of their career?”

3.2 Operationalization of Theoretical Constructs

In this chapter some central constructs are defined, based on the findings of literature review in Chapter II. When operationalizing the thinking model, it became clear to the researcher that the “ideal” thinking model as described in Figure 2, was still too complex. It would make survey’s far too bulky, would introduce a vast number of variables, would take a lot more time than available, in short it would make the research project inexecutable.

To ensure feasibility within the project's time and resource constraints, the model was simplified to be both manageable and falsifiable. This Simplified Thinking Model is shown graphically in Figure 3. More details will be discussed later in this chapter. For

now, an important simplification is that the step of the Impact of CT on the Mental Model of the YP is omitted (and only the mode of CT -internal or external oriented- is measured) and Impact on the Mental Model of others is also defined as Impact. Another simplification is that there is no longer a distinction between the "levels" -group, organization- where we can define Mental Models. Also, the dimension "Conscious/Unconscious" is omitted.

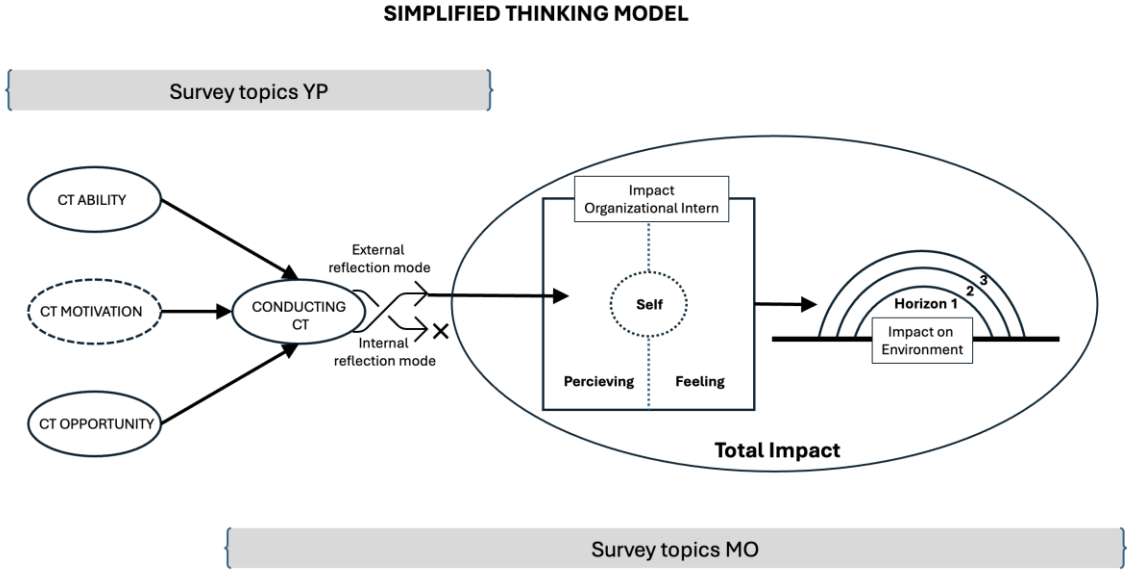


Figure 3:
Simplified Thinking Model

3.2.1 Working Definitions

The definitions in this paragraph are "build up" from other definitions and their explanations, therefore the order is not alphabetically, but follows this "building path". In this research project the next central definitions are used:

- Actor: an entity capable of initiating, conducting, or withholding actions to influence its environment. Actors can be individuals or collective entities (such as organizations, institutions, states, or networks), operating at

varying levels of complexity and agency. They interact within physical, a variety of social and/or ecological systems, shaping outcomes through decisions, behaviors, and strategic actions[2];

- Self: the subjective experience of existence of an Actor, encompassing its awareness of identity, continuity, and agency. It includes a self-concept—a cognitive and affective representation of one's attributes, roles, and relationships—shaped by internal states, interactions, and external influences. The Self enables an Actor to interpret its environment, make decisions, and engage in purposeful action.
- Model: an Actor-driven, purposefully simplified representation of reality that includes selected elements deemed important and the logic of their interactions. It serves to explain past phenomena, explore present dynamics, and estimate future outcomes. Models can be conceptual, mathematical, computational, or physical, and they function as tools for analysis, prediction, and decision-making by abstracting complexity while retaining relevant structure and behavior.
- Mental Model: an conceptual representation of the Self and the reality around it, consisting of -highly simplified- an image of the “Self” and (un)conscious perceiving and feeling (LeDoux, 2020). Mental Models, above all, serve the “Self” to survive and thrive, and therefore to maintain a coherent perception of the Self and the world, even when it does not accurately reflect reality (Cline, 2016).

[2] Where a source is not explicitly given, the definitions are the product of iteratively prompting definitions and wishes as a starting point by the researcher, summarizing and generating alternatives by ChatGPT and manual enrichments and corrections by the Researcher. Starting points for the prompts were manual suggestions of the researcher and the theory from Chapter II.

- Critical Thinking: Critical Thinking is the process of transforming data into information, analyzing information, making inferences, evaluating arguments, and solving problems through applying rigor logic, clear reasoning, intellectual dispositions, and ethical integrity. It involves recognizing and overcoming cognitive biases to develop a Mental Model that accurately reflects reality[3].
- Impact: the effect of an Actor's actions on the external phenomena[4] deemed relevant within its Mental Model. It is inherently relative, shaped by the Actor's Self-concept, values, and worldview, which define the system's boundaries and determine what is considered meaningful change.

N.B.: Measuring Impact involves tracking changes in selected indicators—such as timeframes, effect sizes, directionality, and relevance—relative to the Actor's Mental Model. Given the diversity of Actors and their subjective definitions of significance, Impact should better be understood through the relative shifts in prioritized phenomena rather than through an absolute, universal framework.

3.2.2 Phenomena and Variables

In this research several variables, describing various phenomena, were used. As will be explained in paragraph 3.5, these variables were used in (conducting and analyzing) a Survey. In order to avoid double mentioning, in this paragraph the most relevant variables, phenomena to be measured and the logic behind this, are already given. A detailed description of every used variable and corresponding question in the

[3] The tension between the definitions of CT and Mental Model regarding "reality" is intentional
[4] Impact is, in this study, directed at "to the YP external phenomena", because of the goal of the study. Note that "Impact" in other studies also may encompass effects to the "Self", e.g. "Me using steroids has an Impact on my body". The effect of actions conducted by the YP on the Mental Model(s) of other actors at individual or group level is external to the primary Actor and is therefore also considered as Impact.

survey is given in appendix A: Variables and Survey questions. The CT test will be separately discussed in paragraph 3.7.

Variable: ID

- Phenomenon: Instance of the respondent
- Explanation: part of the used tooling. Used as anonymized linking key in YPID and YPMatchesMO

Variable: YPMatchesMO

- Phenomenon: Used as anonymized linking key for linking the Young Professional and the Manager (or Senior Colleague), in this test further referred to as MO
- Explanation: letting the YPs score their Impact themselves is probably not very accurate. Therefore, the MO is asked to give his impression of the CT conduct, and the Impact the YP makes.

Variable: YPTeamSize

- Phenomenon: the number of colleagues the YP works with on a daily basis.
- Explanation: it is possible that the CT conduct and/or the Impact is affected by the team size. Does it matter if the YP is working alone, in a small team or in a large team?

Variable: YPEduLevel

- Phenomenon: the “highest” education level that the YP completed (University of applied Sciences, Research University of Doctoral)

- Explanation: it is possible that the CT Ability, CT Conduct and/or the Impact is affected by the education level of the YP. Are there differences and if so, what are the differences?

Variable: YPJobMonths

- Phenomenon: the number of months a YP, after completing the HEI-study, is in a "real" working environment.
- Explanation: it is possible that the CT Ability, CT Conduct and/or the Impact is affected by the number of months the YP is in a "real working environment"? Is there a correlation between the number of months in the working environment and the reported CT variables and/or Impact?

Variable: YPEduProfit

- Phenomenon: what is, to the respondent, the result of the HEI-study that is still very profitable to him/her?
- Explanation: in sessions for student information, given by HEI's to help students make a choice which study to choose, (mostly) four main reasons are given. These reasons are having a diploma which is needed for a suitable Job; being able to think on a "HEI"-manner (more or less comparable to CT); gaining knowledge of the field of study and getting a "network" of friends. The choice of the beforementioned dimensions is based on experience of the researcher, so a fifth "free" field is given. Might the respondent(s) think of other profits, this dimensions can be added. How do YPs

feel about the importance of CT (thinking on a HEI-manner)? This question also circumvents a problem that was introduced when the researcher chose to simplify the "Thinking Model" derived in Chapter II for reasons of keeping the research process manageable. When all the levels of the Mental Model, and the effect of every level of Mental Model on every level on the "outer world', would be researched in detail, the survey would become far too bulky. So the choice was made to define "conducting CT" and "Impact" in terms of things the MO sensed. This way we would lose an idea of the effect CT has on the "inner world" of the YP. This question tries to address this missing phenomena.

Variable: YPImpactEnvir

- Phenomenon: the extent to which, according to the YPs, their actions influence their environment.
- Explanation: Impact is, as discussed in Chapter II, an umbrella term. Is it, despite that, possible to simply ask one overarching question about Impact that correlates strongly with the more detailed way it is measured in this study? If so, it would make the before and after training-measurement much easier and faster. Is there a strong correlation between this variable and the detailed measurement?

Variables: YPBelonging, YPCTAllowed and YPCTFreedom

- Phenomenon: to what extend does the YP feel that his voice is a Friendly/supporting voice in the group (whistle-blower theory) and

how much freedom does the YP feel to differ from the opinion in the group?

- Explanation: directly stemming from the theory in Chapter II, it is important to know if the YP feels free to let his voice be heard. If not, I will not invite the YP to the training (to avoid damage). Also, does the score correlate somehow with the conduct of CT? And is the idea of the MO about this subject (the freedom the YP has) the same?

Variable: YPReflectionMode

- Phenomenon: If the YP conducts CT, to what extend is it aimed at the outer world?
- Explanation: based on experience the researcher developed a climbing scale of "direction" of CT behavior, from "not at all" via "clearing the task of the YP"; "offering alternative HOW's to others"; "Letting others think about the WHY" up to "Coaching others to challenge their Mental Model". This scale was tested (and spontaneously recognized by the interviewed persons) during the last test of the survey. If the CT is not aimed at the outer world (not level 3 or up on this five-point scale), and/or the CT is not seen as "Friendly/supporting" (see the description of YPBelonging, YPCTAllowed and YPCTFreedom) there is no "Friendly outwards directed CT". In this research, in some relevant correlation tests, the results of respondents with the scores 1 or 2 are therefore omitted. So, this variable is a "sorting" variable during the number crunch process. This is also the result of simplifying the developed

model in Chapter II to become a less bulky survey and therefore a higher rate of responses.

Variable: YPChallengeScoreReported

- Phenomenon: To what extend is the YP able to conduct Critical Thinking?
- Explanation: a separate CT test, fit for this project, was developed (see paragraph 3.7.2 for more information). Because this developed CT ability test is timed and the survey isn't, the respondents must go to a separate environment and afterwards fill out their score. This variable serves two purposes. The first is to measure to what extend the YP is capable of thinking critically. The second is to see whether the CT is taking this survey serious and is not cheating. It is easy to fill in a higher score then obtained in reality (although this can be detected by the researcher). So, to what extend is the YP capable of thinking critically and has this score a correlation with for example conducting CT and/or making Impact? And, to what extend does the respondent try to give reliable answers.

Variables: MOImpactEnvirYP, MOBelongingYP, MOAllowedYP, MOCTFreedomYP, MOReflectionModeYP

- Phenomena: the measured Phenomenon stem with the Phenomenon described above for the YP. However, this phenomenon is now scored not by the YP, but by the MO. To find

the dexcription of the corresponding variable ~~omit~~ “YP” and *replace* “MO” with “YP” (E.g.: *MO*ImpactEnvir~~YP~~ => YPImpactEnvir)

- Explanation: To what extend do the MO and YP have the same idea about the phenomenon?

Variable: MOBranche, MOISICS, MOOrgCat

- Phenomenon: Information about the environment the YP and MO are working in.
- Explanation: The MOBranche is filled in by hand by the MO. Letting the MO choose between the possible 29 standard possibilities is unworkable. Therefore, the researcher assigns, afterwards, by hand the MOISICS category, a UN standard (*UNSD — ISIC*, no date). This UN list is also the origin of the categorization used by the Dutch Chamber of Commerce and the Dutch Bureau of Statistics (CBS). MOOrgCat is a slightly simplified category the Dutch Chamber of Commerce uses to give an idea about the size in terms of employees and Annual Turnover of a Dutch SME (*Mkb-toets*, no date). These variables are used to find out if there is a correlation between the size of the organization and the Impact.

Variable: MOCTBehaveYPTotal

- Phenomenon: The extend of CT behavior, in the eyes of the MO, the YP shows.

- Explanation: Based on 8 dimensions, where each dimension is two-sided described, the MO scores on a 5-points, the conduct of CT the YP shows. Then this variable is calculated (Figure 4). The number goes from 0 (shows no CT behavior at all) up to 100 (Shows all of the given signs of CT behavior completely). This number is used in multiple tests: does the CT behavior correlate with (variable X)? The 8 dimensions were obtained using the literature review of the researcher and the Creative Commons-work of the University Tilburg introduction to Critical Thinking, written in Dutch, from the Tilburg (Vlerick, 2022). Because of the legal conditions of this document, it was allowed to feed it to ChatGPT as the basis of a prompt. The LLM was asked to use this document and was asked to generate eight covering, non-overlapping but more or less in size equal and weight equal behaviors that could be observed by bystanders. After substantive manual corrections and assuring the language-level was not above a B1 level (*Global scale - Table 1 (CEFR 3.3): Common Reference levels - Common European Framework of Reference for Languages (CEFR) - www.coe.int*, no date) ChatGPT was asked, for each behavior separately, to generate an opposing description in order to get a "Semantic Differential Scale". These descriptions were also manually corrected, checked on the B1-level, non-doubling with the other opposite descriptions and "fit" with the corresponding description. This way chances on respondents having different interpretations of descriptions of behavior was

minimized. The use of eight dimensions, and not seven or nine, was chosen because it "felt" to the researcher like a reasonable pay-off between keeping the survey short enough and still covering the construct.

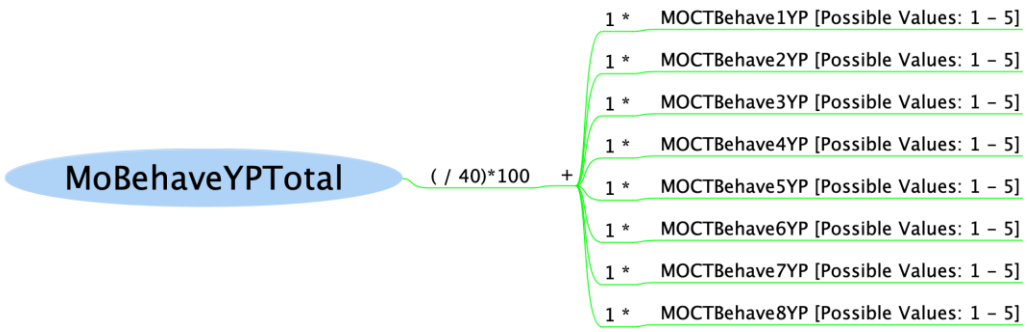


Figure 4:
Calculation of CT Conduct

Variable: MOImpactOrgTotal

- Phenomenon: what is, in the eyes of the MO, the Impact the YP makes on the Mental Model of the organization.
- Explanation: This calculated field (Figure 5) adds the weighted Impact on Feeling, Perceiving and the Self into 1 number, varying from 0 (no Impact) to 5 (maximum Impact). Although the weights are mainly based on experience of the researcher, there is also scientific evidence that it is easier to change feeling and perceiving than changing the "Self" (Polivy, 2001)(Bertsch, 2011). The credibility of the weights was checked against the actual measurements afterwards (see paragraph 4.2.1).

Variable: MOImpactEnvirTotal

- Phenomenon: what is, in the eyes of the MO, the Impact the YP makes on the environment of the organization.
- Explanation: This calculated field (Figure 5) adds the weighted Impact on different "horizons" of change, from the HOW (doing things better/smarter, etc.) to the WHAT (doing different things) and the WHY (reason for existence), a concept familiar in innovation models for a VUCA world, see (Sinha and Sinha, 2020). This variable may vary from 0 (no Impact) to 5 (maximum Impact). Although the weights itself are mainly based on experience of the researcher, there is also scientific evidence that it is easier to change the How, then the WHAT and then the WHY. (Becker and Endenich, 2023). The credibility of the weights was checked against the actual measurements afterwards (see paragraph 4.2.1).

Variable: MOImpactTotal

- Phenomenon: the total Impact made by the YP in the eyes of the MO.
- Explanation: This calculated field (Figure 5) adds the weighted Impact on the Organization and the Impact on the environment. The weights are mainly based on experience of the researcher and in line with for example Covey's Circle of Influence theory (Covey, 2020), suggesting that it is easier to make a change in your

direct environment than in the environments environment. This number may go from 0 up to 19[5]. The credibility of the weights was checked against the actual measurements afterwards (see paragraph 4.2.1).

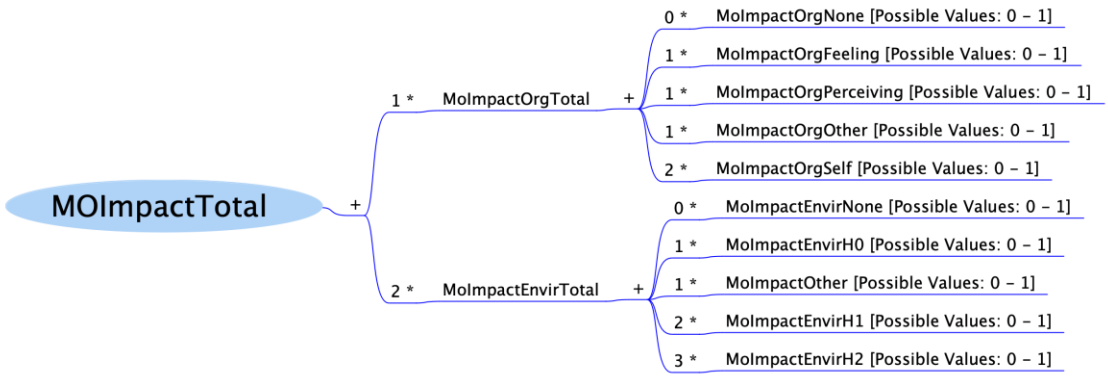


Figure 5:
Calculation of Impact

Variables: MOEduLevel, MOEduProfit

- Phenomenon: the measured phenomena stem with the phenomena described for the YP earlier. To find the corresponding variable and replace "MO" with "YP" (E.g.: MOEduLevel => YPEduLevel)
- Explanation: it matches the description and logic given earlier, however, this time it is about the benefits the MO thinks his own education gives.

[5] This Impact-score is not "normalized" in a scale from 1 to a 100, unlike the MOBehaveYPTotal. Without wanting to start a philosophical discourse, the researcher felt that Impact may go beyond the for this research relevant defined scale, e.g. Impact at a "spiritual" level. CT behavior has more of an agreed "maximum" and therefore is "normalized". Impact has (outside the model described in this research) probably no "maximum" and is therefore not "normalized". This choice is of course arguable but is has no effect on the conclusions that can be drawn based on the analyzing process.

3.3 Research Purpose and Questions

The purpose of this research is to find out if CT, conducted by HEI-graduated YPs, really is of importance for them making Impact at the start of their careers. If it is not, lots of time can be spared during their HEI-education. If it is, students, HEIs and the society can rest assured that the time and money spend on mastering CT is well spend. Insights probably lead to suggestions for students, YPs, HEIs and Employers to enlarge the Impact YP's can make.

As stated in paragraph 1.5 the first two main questions are "Is there a relationship between Critical Thinking by HEI graduated Young Professionals and to the Impact they make at the start of their careers?". And, if so, "Does Critical Thinking by HEI graduated Young Professionals lead to making Impact at the start of their careers?". In paragraph 3.4 these research questions are broken down into questions that can be tested with hypotheses, based on measurements of the variables described in paragraph 3.2. Insights in the measured variables can also give input to answering the third question "What actions can YPs, HEIs and Employers take regarding learning and conducting CT that will most probably lead to the Impact YPs make at the start of their career?"

3.4 Breakdown of the Research Questions

In this paragraph the questions, as described in paragraph 3.3, are broken down into questions that, in turn, can be broken down into hypothesis that can be falsified once the measurements are done. It may be noticed that in paragraph 3.2.2 already an

important research question is already partly answered: how to make the phenomena considered in this research measurable?

3.4.1 Usability of the Thinking Model: Calculated Variables

The calculated variables are based on assumptions, stemming from the thinking model developed in Chapter II. So ideally the real measurements should not conflict with the expectations based on this thinking model. Or, even better, predictions stemming from the model about the outcomes should come true. Otherwise, the outcome of the calculations, as described in paragraph 3.2.2, are useless even before they are made. So, the question is "Does the thinking model, regarding the assumptions behind the calculated Impact score, has any explaining or prediction power?". In the thinking model assumptions are made about some phenomena being easier to accomplish than others. The chosen "weights" in the model are based on this idea. So, looking at the data, the number of times the "easier to accomplish" phenomenon is reported should exceed the number of times the "harder to reach phenomenon" is reported. If not, this assumptions behind the thinking model are questionable. Based on this thought and the thinking model, the next predictions can be made:

1. In the Mental Model, it is more difficult to make Impact on the "Self" than Impact on the feeling and perceiving. The number of reported scores on both MOImpactOrgFeeling and MOImpactOrgPerceiving should individually exceed the number of Impact scores on MOImpactOrgSelf.
2. It is more difficult to make Impact externally than internally. So, the number of reported scores on the Mental Model as a whole

(MOImpactOrgTotal), should exceed the number of reported scores on the External Impact (MOImpactEnvirTotal). N.B.: notice that both variables consist of 3 sub variables, which makes "normalization" unnecessary.

3. In the external Impact, the "lower levels" are easier to influence than the "higher levels". So, the number of reported scores of Impact on a lower level should exceed the number of reported scores on the higher levels (#scores on MOImpactEnvirHa > #scores on MOImpactEnvirHb, where a < b)
4. According to the thinking model, Impact on the Mental Model proceeds Impact extern. So, it is possible to make Impact on the Mental Model without making Impact extern, but it is mandatory to make Impact on the Mental Model in order to make Impact in the outside world. So, there should be no cases where MOImpactEnvirTotal is scored, and MOImpactOrgTotal isn't. But it is possible that there are cases where MOImpactOrgTotal is scored, but MOImpactEnvirTotal isn't.

(N.B.: It should be mentioned that the above questions are a quite rough, quick and dirty way to invalidate a model. Validating models is an art on itself and is usually done with a statistical technique called factor analyses, in this case confirmatory factor analysis. However, this technique requires a vast amount of data, with at least 100 "observations", which will not be available during this research (RC MacCallum *et al.*, 1999))

If the above hypotheses come true, it would also be valuable to check whether the chosen weights are reasonable. Therefore, in paragraph 4.2.1 the hypotheses weights and formula to reach an overarching "Impact" score are checked against the reality.

Differences between the data and the hypothesized weights should not be (statistically) significant.

3.4.2 The relationship between CT and Impact

The search for a relationship between CT and Impact can be divided in two parts: correlation and causation. For the correlation part, the next questions can be formulated:

1. Does CT OMA (the combination Ability, Opportunity and Motivation) correlate with Impact?
2. Does CT Conduct correlate with Impact?

For the causation part ("caused by"), the above questions can be repeated, but with the words "correlate with" replaced with "lead to". As the model of Chapter II is further simplified, the question "Does CT Conduct lead to changes in the Mental Model of the YP" and "Do changes in the Mental Model of the YP lead to Changes in the Mental Model of the Group/Organization" is omitted. Changes in the Mental Model outside the YP are defined is Impact (see the definitions in Paragraph 3.2.1).

3.4.3 "Miscellaneous constructs", CT and Impact

In Paragraph 3.2 some Miscellaneous constructs are chosen, that could probably relate with CT an Impact. Questions formulated for this relationship:

1. Do (combinations of) the Miscellaneous constructs have a correlation with CT OMA (the combination Ability, Opportunity and Motivation)?
2. Do (combinations of) the Miscellaneous constructs have a correlation with CT Conduct?

3. Do (combinations of) the Miscellaneous constructs have a correlation with Impact?

\

3.4.4 Logic of the Research Questions summarized

In order to be able to answer the research question in Chapter V more easily, the research questions can be structured as follows:

Main questions:

- Is there a relationship between Critical Thinking by HEI graduated Young Professionals and to the Impact they make at the start of their careers?
- Does Critical Thinking by HEI graduated Young Professionals lead to making Impact at the start of their careers?
- What actions can HEIs, Employers and YPs take regarding learning and conducting CT that will most probably lead to the Impact YPs make at the start of their career?"

In order to be able to answer the main questions, the next sub questions were formulated:

- Using the Thinking Model developed in Chapter II, how can CT OMA (Ability, Motivation, Opportunity), CT Conduct, and Impact on the Mental Model and Environment be operationalized?
- Once the measurements are done, did the operationalization of the Thinking Model "hold"?

- On the premise that the model holds, what can be said about the correlation between the defined variables and Impact?

3.5 Research Design

This project initially encompassed a "mixed method" approach: a combination of a Survey and an experiment in the tradition of Action Research. With the survey the researcher aimed at finding out if there is a correlation between conducting CT by YPs and them making Impact.

With the Action Research the researcher aimed to find out if supporting conducting CT by YPs in their working environment creates even more Impact, and therefore if there is a short-term cause-effect relationship between conducting CT and making Impact. This support would take place on all of the three "OMA" elements mentioned in the Simplified Thinking Model (Figure 3): the Opportunity to conduct CT, the Motivation to conduct CT and the Ability to conduct CT.

In order to create maximum validity of the findings, this cause-effect part would ideally use a "clean" double-blind research design as would be done in a genuine Experiment. However, because the research would be conducted in the real world and not in a laboratory, this would raise both ethical and practical issues. Therefore, this project would use the results of the Survey as a "Baseline" for interpreting the results of the Action Research part. To be able to do that, the results of the Action Research part would (also) be measured with the same tools as were used in the survey part. The participants would be drawn randomly from the survey respondents. In order to minimize possible negative effects for the participants in the Action Research part, the participants would only be chosen from organizations where the manager interprets the YP as having a neutral or "supporting voice" (as known from the whistleblower research). To minimize

the risk of "spreading out" negative effects in case of unforeseen events, the Action Research would take place on three separate small groups of five participants. With the forming of these groups, possible competition considerations between the employers would be taken into account. Participants of the Action Research groups would receive some "Certificate of Participation" at the end. This "reward" is communicated beforehand, as it should (hopefully) increase the number of respondents and participants. However, because of the limited number of respondents that could join this cause-effect part, this part of the research could eventually not be meaningful conducted within the given timeframe.

Therefore, the conclusions of this research project are solely built on a non-experimental, correlational study, which was essentially a survey-research design.

3.6 Boundary Conditions

This research project had on top of the "usual" conditions like Ethical considerations, some conditions and considerations that were of solidly influence. The following were recognized:

- GDPR. Tools used should not bring personal data out of the EU domain and the researcher should be able to ensure confidentiality.
- Use of AI. The possible use of AI for processing the data should be clear to the respondents and also not violate the GDPR. In Paragraph 3.10.3.1 the text is given where the respondents agree upon, regarding the use of AI.
- Limited time. The time for the research project is as a whole is limited to two years. As we will see in paragraph 3.9 this boundary condition withheld the researcher from the possibility to choose an alternative

timeframe, when the “experiment” part of the research-project as described in paragraph 3.5 could not be conducted because of the limited number of YPs willing to join.

- Limited budget. The out-of-pocket costs for the project should not exceed 2.000 euro.
- Complexity of the language used in the survey. The population structure of the Rotterdam municipality is a-typical compared to the other four big cities in the Netherlands: there are proportionately more lower educated and fewer highly educated citizens, and proportionately less citizens with a western background. Also the average incomes are lower. (*Gemeente Rotterdam in cijfers en grafieken (bijgewerkt 2024!)*, 2024). Universities of Applied Sciences in the Netherlands (HBO-instellingen) generally have a regional function: about 80% of the graduates find their first job in the region (*Versterken van de regio kan niet zonder duurzaam investeren in regionale spilfunctie hogescholen*, no date). Thus, it is expected that many students will find employment in the region. Therefore, the research was conducted in Dutch. When questioning managers, it was taken into account that the Surveys also connect with "self-made" non-university level educated people. To keep the response rate as high as possible in number and quality, concise Surveys in clear formulated terms were preferable to substantively correct, extensive and detailed Surveys. The Survey had to consist of clearly stipulated "need-to-haves" only. This limited the number of variables that could be taken into account in this research project. This focus limited the possibility to find “unexpected gems” in the dataset but increased the validity of the study. As the YPs

have an (applied) university level, the restrictions on the language level for them would not have to be as stringent.

- Time available by YPs and MOs. To maximize the respondent rate, the Survey should be workable. To be concrete, the researcher chose 15 minutes for filling in the survey for both YP and MO as the maximum time the survey should take, and 20 minutes as the maximum time the CT test should take. Longer times lead to fewer respondents and -as people tend to "want to finally get it over and done"-more uniform scores for questions at the end of the form (GALESIC and BOSNJAK, 2009). These choices were "validated" in the testing rounds by asking the testers in person what they would feel as "acceptable". A much shorter timespan as defined was also seen as undesirable, as - citing an interviewed person who took the test in the validating phase- "it could give the respondents the idea that their reactions would not be taken seriously". This boundary led to a simplification of the thinking model as developed in Chapter II. Impact is defined in terms of "everything outside the Mental Model of the YP", and is not measured separately on the personal, group and organizational level. Also conducting CT is narrowed down to actions that the YP challenges the Mental Models of others in a "Supporting/Friendly Voice". See paragraph 3.2 for more information. To keep the time for the respondents as short as possible, there were two paths through the Survey: one for the YP and one for the MO. This "breaking up" limits the possibility of cross-checks but increases the (researchers' expectation of) the amount and quality of reactions.

- Avoiding using the Dutch academic translation of the phrase “Critical Thinking”. Critical Thinking, in Dutch translated as “Kritisch denken”, may have a negative connotation. The word ‘Kritiek” has many synonyms in Dutch, like “snibbig, beschimping, veroordeling, afkeuring, commentaar, bedreigend” (*Synoniemen van kritiek; ander woord voor kritiek*, no date). Translated with Deepl.com, it says in English “snide, taunt, condemnation, disapproval, comment, threatening”. Given the beforementioned arguments about MOs, it was likely that at least some part of the respondents would not be acquainted with the meaning “Critical Thinking” has within the scientific/educational field. In order to improve the willingness of managers/organizations to support CT and accept the results, other labels were tested, and the alternative “HEI-thinking” was used.
- The use of non-standard tests and definitions. Although the developed model is substantiated with a theoretical framework and choices are argued, it is still a model. This means an uncertainty is introduced about the “fit” between the model and real life. Therefore various “safeguards” were built in to find out if the model and reality are inconsistent.
    - Where possible, parts of existing tests were used.
    - As the Survey should be as clear and short as possible, the researcher could not afford to use a lot of “control questions”. Materials were tested beforehand in a limited group, and persons were asked to give their opinion in person. Questions that were clearly interpreted in another way as the researcher intended, or were not understood at all, were rephrased.

- This way of work was also used to find out if the logic behind the answers of respondents corresponded (enough) with the phenomenon the researcher wanted to measure with each question. If, during the interview a relevant phenomenon would be encountered that was not included in the Survey (or the Thinking Model) extra questions could be added. Looking backwards, this did not take place.
- Also, at some critical points the possibility to fill in "free text" fields in the Survey was given; if there could be an answering-possibility or disposition that was not covered by the answers, the respondent could add his/her answer. This was indirect also a measurement of quality: when this "fill in the blank answer" possibility was used only sparsely, the "universe of answers" covered the reality sufficiently enough. As we will see in the paragraph where the results are given, this was the case.
- The quality of the CT test was also measured with Cronbach's Alpha (see Paragraphs 3.7.2 and 4.2.2).

3.7 Instrumentation

In this paragraph the instrumentation that was used during this project is discussed. Also, the tools that were considered, but in the end were not used, will be discussed.

3.7.1 Survey Tools

A search for appropriate survey tool was conducted. At first, within the boundaries as mentioned in paragraph 3.6 in mind, a wide range of alternatives were compared. Because the Rotterdam University of Applied Sciences (RUAS), homebase of the researcher, has its own survey-tool, this one was amongst them. However, this RUAS tool was not chosen for several reasons. The first main reason was the most likely extra coordination between the researcher and the tool-administrator which would possibly cost extra time but would not give extra benefits. The second reason was the unfamiliar environment for the respondents, which could lead to both a lower response-rate and possible technical problems between the respondent's environment and the RUAS-tool (firewalls, limitations of web browsers).

The RUAS also has a contract with Microsoft, making MS-Forms a possible candidate for the survey. Confidentiality can be guaranteed that way, as can the GDPR compliancy. Also, because of the broad use and acceptance of the Microsoft software, also in Dutch SMEs, technical problems between the MS-Forms cloud-environment and the respondents' environment are most unlikely, and unfamiliarity with the tool would probably not be a problem. However, MS-Forms uses a data format that makes it nearly impossible to do statistical tests on the raw data. During this project a "translator" was built by the researcher using MS-Excel to overcome this data-format problem (see Appendix B for a small example). Also, the variety of possible types of questions is, within MS-Forms, quite limited. For example, the possibility of using a "Semantic Differential Scale" is not given. With a (in real life tested) workaround for the standard "Likert score" possibility the researcher found a workable solution for this omission (Figure 6). Also, the possibility of needed "individually timed questions" is not given in Forms. The researcher also found (and tested in real life) a workaround for this problem. See for this solution the introduction text in the reasoning challenge, paragraph 3.7.2.

Other survey tools, also paid ones, were considered, but did not meet the requirements. As said, during the design process the researcher encountered several problems where MS-Forms did not offer a standard solution. This triggered most of the times a new round of searching for a survey tool that did offer the needed solutions ‘out of the box”. However, with the developed workarounds, MS-Forms was at the end the tool that was used for the survey.

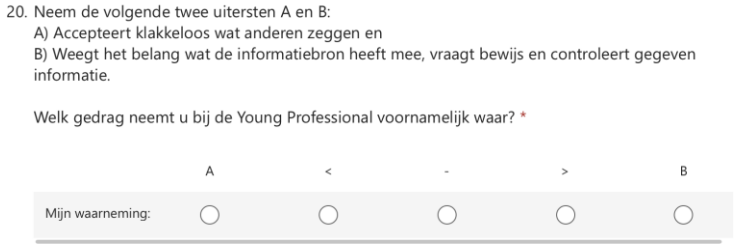


Figure 6:
Likert Scale transformed to a Semantic Differential Scale

3.7.2 Measuring CT Ability

Measuring the Critical Thinking ability and motivation of Young Professionals can be done by a standardized test, as this increases the possibility of generalizing the results of the study at the end. However, there are different theoretical viewpoints about using standard tests compared to tests specially tailored to the situation (Heijltjes, no date).

The researcher considered several options, one of them being a tested Dutch translation of the “Halpern Critical Thinking Assessment” (De Bie, Wilhelm and Van Der Meij, 2015). However, after contacting the authors, it was clear this project was closed, and this survey was not available free of rights and costs. Also, the definitions of CT did not stem with the definition in this survey. More precise, the recognition of

"Thinking Biases" was not covered (Hatcher, 2013). Finally, the results of "standardized tools" might be sound, but at the cost of a lot of time for the respondents, thus violating one of the boundaries as described in paragraph 3.6.

It was concluded that there was no lightweight scientifically "approved" standardized test free of costs, and rights, encompassing the definition of CT in this survey. The conclusion was that a test had to be developed by the researcher himself. The development of this CT test took four rounds.

In the first round the basis was realized. Three questions were used from the publicly available test material website of Fibonnicci (*Home*, no date), aimed at preparing oneself for taking the Halpern Critical Thinking Assessment. The first version of the CT test encompassed 13 Multiple Choice and 4 open questions, where in total 100 points could be scored:

- The first 13 questions were, at 4 points per right answer (52 points in total), aimed at "reasoning". The first 6 questions were pairs of "the same kind of reasoning", the first one of the two being easy and the last being more difficult. Pairs were verbal reasoning, syllogisms, mathematical series, Figure series, and three-dimensional reasoning. The last questions were "type 2 thinking" challenges, amongst them exponential/statistical reasoning.
- The last 4 open questions were four short case-descriptions, where someone suffered from -at least- three thinking biases per case. The respondent was asked to discover if, and if so what kind of, thinking bias could be recognized in the case and to describe the biases. Every correct

bias would lead to 4 points (giving a possible total of 4 questions * 3 biases * 4 points = 48 points).

These cases were generated with OpenAI's GPT-4, when prompted (in Dutch) for "A short work-related case, appealing to Young Professionals, where the leading Figure suffers from [3 biases were given by the researcher, for example: confirmation bias, negativity bias, avoidance of negative information]." The biases were chosen from the lists of well-known biases as given in the work of Vlerick (2022) and additional sources(Cohen, 2015)(Kahneman, 2013)('List of cognitive biases', 2025). Then these AI-generated cases were manually optimized. After optimization GPT-4 was asked to generate an explanation of the biases the leading figure made, in order to be able to give this as a response to the person making the test. These explanations were also manually optimized.

The next open-question cases (originally in Dutch, translation by Deepl.com) were given:

Case A:

A young project manager at a consultancy firm has created a tool in Excel and gets positive feedback on it from a close few colleagues during interim test versions. He ignores comments from other team members because he is convinced of the success of his tool. Just before the tool goes live, he avoids discussion with his boss about possible risks, because he fears the consequences of negative feedback. The tool is applied in a real environment for the first time, and the first negative comment from a customer employee

hits him much harder than all the other positive feedback from employees there. In the end, the tool is not added to the consulting company's tooling set.

Recognizable Biases:

Confirmation bias: The project manager ignores comments from other team members that may contain criticism or suggestions because they do not match his belief that the tool will be successful. He mainly seeks confirmation of his own ideas and excludes other information that might undermine this belief.

Anchoring bias: The project manager may be strongly influenced by the initial positive feedback he has received from a few close colleagues. This early affirmation may form an anchor point, making him less open to later, possibly more critical feedback. Impact bias: The project manager perceives a customer's first negative comment as much heavier than the previous positive feedback. This is an example of impact bias, where the emotional impact of negative feedback is overestimated compared to positive feedback.

Negativity bias: Negative information affects the project manager more than positive information, which is an example of negativity bias. He focuses on the customer's negative feedback, causing him to doubt the usefulness of the tool, despite previous positive feedback.

Avoidance of negative information: The project manager avoids an open discussion with his boss about possible risks, probably because he fears negative feedback. This avoidance behavior can lead to poor risk assessment and preparation.

Overconfidence bias: The project manager seems to have excessive confidence in the quality and success of his tool, based on limited positive

feedback. This overconfidence may cause him to insufficiently consider potential problems or criticism.

Case B:

A young sales manager is working on an important deal. He bases his strategy on the last successful sale he made, because he still has the details of it well in his memory. Unfortunately, this second deal fails. He tells his colleagues the unpleasant news, saying that he could have predicted from the start that this customer would not bite.A third deal he tackles the same, and there too success fails to materialise. He himself is not really bothered by the lack of success, believing that a fourth deal will certainly succeed, because he thinks that after two failures, success is inevitable.

Recognizable Biases:

Availability heuristic: He bases his strategy for the second and third deal on the last successful sale that is still fresh in his mind, without considering the differences between the customers or circumstances.

Hindsight bias: After the failure of the second deal, he claims he could predict from the start that the customer would not bite, while he probably only developed this knowledge after the fact.

Gambler's fallacy: The sales manager believes that a fourth deal will definitely succeed because he thinks that after two failures, success is inevitable, as if there should be a natural balance between successes and failures.

Confirmation bias: He continues to adopt the same approach despite failed deals because he seeks proof that his strategy works, ignoring information that proves otherwise.

Optimism bias: The manager has an overly optimistic view of the outcome of future deals, despite repeated setbacks, which clouds his assessment of risks and opportunities.

Case C:

A marketing analyst, new to the business, sets a budget for a new campaign. In the absence of comparables, he asks the client what he thinks such a campaign will cost and bases his budget estimate on the amount mentioned by the client. He comes to an agreement with the client and gets the job. In putting together the team that will set up the campaign, the marketing analyst looks for a team member who can bring diligence and quality awareness to the team. He asks for a colleague whom he has noticed always expresses himself very carefully in words, because he assumes that this colleague will then probably also be careful in his project work. He has such a good feeling about his approach to the campaign that, when the client asks him to make a prediction about the results he promises results that will be hard to achieve.

Recognizable Biases:

Anchoring bias: The marketing analyst bases his budget estimate on the amount mentioned by the customer. This is an example of the anchoring bias, where the first piece of information (in this case, the amount mentioned

by the customer) is used as a reference point, even if it is not based on good information.

Representativeness heuristic: The analyst assumes that the colleague who expresses himself carefully in words will also be careful in his project work. This is an example of the representativeness heuristic, where one characteristic (careful word choice) is wrongly seen as representative of another characteristic (diligence in project work).

Overconfidence bias: The marketing analyst is overconfident about his approach and promises results that are difficult to achieve. This is an example of overconfidence bias, where someone is too confident in their own ability or the outcome of their decisions.

Case D:

A young entrepreneur decides to launch a startup after following several influencers with success stories of companies that started from a garage. He is betting on making a popular tech trend of value because many other startups are doing the same, although it is still a bit of a search how relevant it is to the market he has experience in. His startup is struggling and that worries him, but he keeps going because he has already put so much effort and money into it. In the end, his startup appears to be making a living.

Recognizable Biases:

Bandwagon effect (Me too thinking): The entrepreneur bets on a popular tech trend simply because many other startups are also doing it. This is an example of the bandwagon effect, where someone tends to do something

because others are doing it too, without critically considering their own situation or relevance to the market in which they operate.

Survivorship bias: The entrepreneur bases his decision to launch a startup on success stories of companies that started from a garage, as told by influencers. In doing so, he ignores the many stories of startups that were not successful, which is an example of survivorship bias. This is the tendency to look only at successful examples and forget that there are many more failures that are not in the spotlight.

Sunk cost fallacy: The entrepreneur keeps going with his startup because he has already put so much effort and money into it, even when the company is struggling. This is an example of the sunk cost fallacy, where someone decides to continue with an investment because a lot has already been invested in it, without looking at the current and future costs and benefits.

This first version of the test had a try-out under the direct colleagues of the researcher and was very insightful (a label that might be considered as a euphemism for "disaster").

In order to be able to score the last 4 questions automatically, the researcher developed a prompt for ChatGPT that scored the answers automatically. This scoring of answers on the open questions by the LLM led to impressive results. However, this way of evaluating the answers was eventually abandoned. Not in the first place because LLM's evaluate texts with a certain degree of randomness (*Why does the answer vary for the same question asked multiple times - Community*, 2024), so with every run measures should be taken to make sure the evaluation is still acceptable. This way the costs of

evaluating the answers given by the LLM would exceed the benefits. Also, the evaluation could not be integrated in MS-Forms, so the respondent would have to wait for his/her score, probably leading to lower respondent rates. But the most severe problem was a social one: because respondents wanted to "score high or die", they stuck with difficult questions and the total test took far too long to complete. The quality of the answers on the open questions also diminished (open question 1 was answered in more detail than open question 4). And finally, one of the respondents used ChatGPT herself to come to the "right" answers on the open questions, which made this part of the test results useless. So, the learning points for the researcher were threefold: instantly give the "score" back at the end of the test, use timed questions and do not use an open question format.

The second round these problems were "solved" in a new version of the CT ability test. Some questions were altered, and the open questions were replaced. In this version the test still encompassed 13 Multiple Choice, and 12 "yes/no" questions, where in total 100 points could be scored:

- The first 13 questions were, at 4 points per right answer (52 points in total), aimed at "reasoning". The first 6 questions were pairs of "the same kind of reasoning", the first one of the two being easy and the last being more difficult. Pairs were verbal reasoning, syllogisms, mathematical series, Figure series, and three-dimensional reasoning. The last questions were "type 2 thinking" challenges, amongst them exponential/statistical reasoning. Changes were made based on the results of the first tests. The final result can be found later in this paragraph.
- The last 12 questions were based on the same case and thus formed a "storyline". This story was manually built by the researcher, taking one of

the cases of the first version as a starting point. Every question described, in a maximum of 2 sentences, the “leading Figure” undertaking an action. That action could be based on a thinking bias, or on clear reasoning. The respondent was asked whether the person suffered from a thinking bias or not. With every question worth 4 points, a total of 48 points could be scored.

- A timer on the test as a total was introduced. It was, within MS-Forms, impossible to set a timer on each question individually. So, the researcher wanted to test if an alternative solution was clear to the respondents. At the start of the test the following instruction was given (originally in Dutch, the next translation conducted with deepl.com): ”Under each question, in light gray, are 2 numbers in square brackets. The first number indicates the maximum time available for that question. The second number indicates the minimum remaining time required for the entire Survey. The top of the screen shows how much residual time you actually have left. So make sure that the time at the top of the screen remains larger than the minimum required residual time. You may decide how much time you spend on an answer, but cannot take more than 20 minutes over the entire test. You will not receive points for unanswered questions. So don't get stuck on a difficult question and pick some...wrong answers don't earn you minus points. Answering quickly if you know it is convenient; if you have time left at the end, you can always jump back to a question you found difficult.”

This second version was tested amongst 20 students who volunteered. They all started at the same time. The winner (most points, and on a draw the fastest player)

earned €20,-, number 2 and three €10,- . During the test the researcher polled using a minute/question matrix, waking around, which question was answered by the volunteers at every minute. Thus, afterwards a better estimated guess of the time each question took could be calculated, which was used in the optimized third version. Analysis of the results showed that, with "weak questions" omitted, the Cronbach Alpha could go from a 0,711 up to 0,766, which was a good enough test reliability (Tavakol and Dennick, 2011). There was a slight negative, not statistically significant correlation between the time a respondent took to conduct the survey and his/her result, and every respondent managed to complete the test in time, so the given timeframe to complete the test was sufficient enough. Asked afterwards, the respondents declared that the instructions regarding the timer were clear to them. Based on this test, some questions were rephrased or completely changed, and it was concluded that this test could be completed in a time under 20 minutes.

The third version of the test was conducted amongst 40 students, also volunteers. The Cronbach Alpha was 0,72. The mean was 56,3 points. The distribution of the total scores was as expected, compared to "normal exams" the researcher uses in his daily work; there was, visually, a "Normal Distribution enough" curve and also the Shapiro Wilk P was 20.6% (which is far above the required 5%). Therefore, we may assume the data stemming from this test was normally distributed.

So, given the prerequisites are met, it was possible to do a T-test against the score that would be obtained if the results were based on luck. If the obtained points by the respondents were the result of pure "guessing", the mean of the number of points would be around (13*0,25 chance +12*0,5 chance)=3,25+6=9.25 correct answers. Which would result in 9,25 times 4 points = 37 points.

A one-tailed one-sample T-test was conducted to find out if there is a significant difference between the Mean of the total score and 37. The results showed a significant difference, $t(39) = 7.51$, $p < 0.001$, $d = 1.19$. The average score Total score was 56.3, $SD = 16.3$. This indicates that "something" is measured with a good enough test reliability. This "something" will most likely be CT, because the questions were obtained from this construct. So, the CT test was accepted by the researched as a "good enough" tool to measure the CT ability conform the definition as given in this research paper.

The fourth and final test was more technical and aimed at the process: was it possible to link in the "overarching" survey to the CT ability test, and let the respondent fill out the score in the original survey once he or she was finished with the CT ability test? And was it not experienced as annoying? This test was done with a "real" YP and MO and led to a satisfying enough result.

The translated version of the CT Ability test (translation using Deepl.com and manually corrected), including some remarks, is given below. To be able to differ the translation of the text of the CT Ability from remarks made in this thesis, the translation is cursive. Questions that are mandatory are marked with "*". Remarks from the researcher, solely for the readers of this report, are placed between accolades "{" and "}"

Reasoning challenge (100 points)

Answer the questions and see how strong you are at pure and logical reasoning!
This challenge consists of 2 parts and takes a maximum of 20 minutes in total. Part 1 consists of tasks in which you can demonstrate your logical thinking. Part 2 consists of recognizing thinking errors. In total, there are 25 questions.

Under each question, in light grey, are 2 numbers in square brackets. The first number indicates the maximum time available for that question. The second number indicates the minimum residual time required for the whole Survey. The top of the screen shows how much residual time you actually have left. So, make sure that the time at the top of the screen remains larger than the minimum required residual time.

You can decide how much time you spend on an answer, but you cannot take more than 20 minutes for the whole test. You will not get points for unanswered questions. So don't get stuck on a difficult question and pick some...wrong answers don't earn you minus points. Answering quickly if you know it is convenient; if you have time left at the end, you can always jump back to a question you found difficult.

First, you are asked to enter a 'Search Key'. This is the email address you use for these surveys. The content of this input field is stored in the analysis set as an encrypted Hashcode and thus anonymized. This hash code allows us to match your parts of the surveys.

By participating in this test, you agree to the research conditions (see https://edu.nl/fypvv).

You will receive the results immediately after submitting the test.

Good luck with the puzzle!

PART 1: Logical reasoning

Please enter the search key (your e-mail address) first. This is followed by 13 multiple-choice questions. Choose the most correct answer each time!

***1** The search key: which email address may we use to match your survey items*
(NB: this field will be anonymised in the analysis set)

{text field with emailadress validity check}

2 *Take this statement to be true: 'Everyone who works receives money '. So which proposition is 'true'? **

[Time: 1/20] (4 points)

A Everyone who receives money works

B Everyone who does not work does not receive money

C Everyone who does not receive money does not work

D All previous statements follow from the first statement and are 'true'

3 *Take the following statements to be true' "All students are smart people " and 'Some smart people are rich '. Then you may infer... **

[Time 1/19] (4 points)

A All students are rich

B Some students are rich

C All rich people study

D None of the above

4 *'Listen' stands to 'Silent' as 'Night' stands to... **

[Time: 0.5/18] (4 points)

A Rest

B Thing

C Dark

D Invisible

{This challenge was not translated, but adapted to English)

5 *[1] stands to Aristocrat as Sewer stands to [2]... **

[Time: 1/17.5] (4 points)

A [1] = Castle ; [2] = Rat

B [1] = Refinement ; [2] = Stench

C [1] = Millionaire ; [2] = Poor

D [1] = Street ; [2] = Castle

6 *Complete the sequence: 1 ; 3 ; 4 ; 7 ; 11 ; ? ... **

[Time: 0.5/16.5] (4 points)

A 13

B 15

C 18

D 19

7 *Complete the sequence: -1 ; 14 ; 1 ; 11 ; 3 ; 15 ; ?... **

[Time: 1/16] (4 points)

A 5

B 1

C 13

D 7

8 *On day 1, duckweed falls into a clean pond. Every day the amount of duckweed doubles, until exactly after 2 weeks (14 days) the whole pond is covered with duckweed. On which day is the pond half covered with duckweed? **

[Time: 0.5/15] (4 points)

A the 7th day

B the 9th day

C the 11th day

D the 13th day

***9** A croissant with cheese costs €1.10. Cheese is one euro cheaper than a croissant. How much does the cheese? **

[Time: 0.5/14.5] (4 points)

A € 0,05

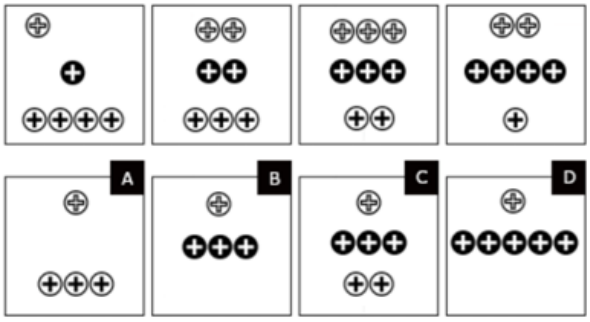
B € 0,10

C € 0,15

D € 0,20

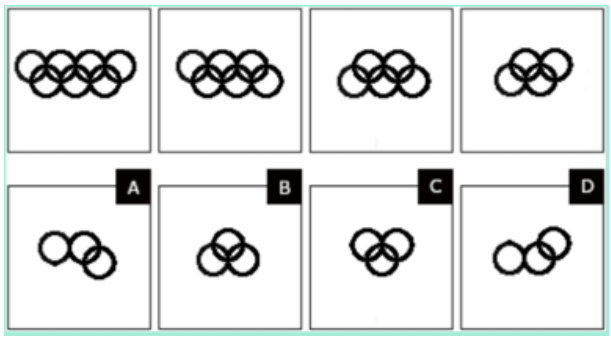
***10** Which Figure follows the top four Figures? **

[Time: 0.5/14] (4 points)



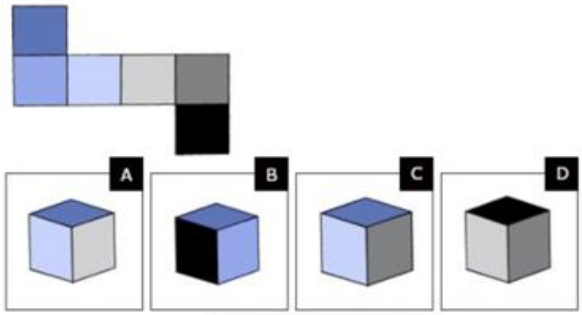
***11** Which Figure follows the top four Figures? **

[Time: 0.5/13,5] (4 points)



12 *Which Figure would be created if you folded this Figure into a cube? **

[Time: 1/13] (4 points)



{As stated, Question 10, 11 and twelve were obtained from https://www.fibonicci.com)

13 *In a lottery, a ball is drawn 7 times in a game round from a bin with 100 balls. Each ball is put back into the bin immediately after that one ball is drawn. The balls in the bin are the numbers 0 to 99, all numbers occur and each ball has a different number on it. You get to choose your own set of numbers. Which of the following strategies offers, when you play multiple times, the least chance of winning? **

[Time: 1.5/12] (4 points)

A Play each game round with ANOTHER set of your own randomly made-up, different, numbers (Example: 1,3,45,54,60,78,79)

B Continue to play every game round with THE SAME set of, mutually similar, numbers (Example: 33,33,33,33,33,33,33,33)

C Each round, play with the 7 numbers that have been MOST frequently called by the game management up to that point.

D All strategies listed here have EXACTLY the same chance of winning or losing

***14** A customs officer at an airport searches people. In 80% of cases, for passengers to be searched who are smuggling, he correctly senses in advance that this passenger is smuggling. In what % of cases, then, for passengers to be searched who are NOT smuggling, does the customs officer correctly sense in advance that this passenger is NOT smuggling? Choose the most correct answer...*

[Time: 1.5/10.5] (4 points)

A 20%

B 60%

C 80%

D This cannot be determined from this data.

{N.B. as the first question was used to get the email address from the respondent, the first 13 questions were numbered 2 up to 14}

PART 2: Recognizing Thinking Biases

Here is a case of a young entrepreneur. This case is described in parts. For each part, the young entrepreneur made a 'thinking fallacy' (in English: 'Cognitive Bias'). A thinking fallacy is incorrect reasoning, causing someone to draw non-rational, illogical conclusions. For each question, indicate whether the entrepreneur's thinking fallacy occurred. Each correct answer earns points. Good luck!

[Reading time: 1/9]

15 *A young entrepreneur decides to launch a startup after following several influencers with success stories of companies that started from a garage. Inspiring, that's how it should be!**

[Time: 0.5/8] (4 points)

A He makes a fallacy

B He does not make a fallacy

16 *He is betting on making a popular tech trend of value because many other startups are also doing this, although it remains to be seen how relevant this is to the market he has experience in. After all, good example follows good practice. **

[Time: 0.5/7.5] (4 points)

A He makes a fallacy

B He does not make a fallacy

17 *He further explores the market by reading bank reports, trade literature and spelling out trend analysis, such as data from the Central Statistical Office. To get ideas about needs in this market, he talks to many different people familiar with his market. **

[Time: 0.5/7] (4 points)

A He makes a fallacy

B He does not make a fallacy

18 *He comes up with several ideas, clearly having 1 favorite idea. He decides to develop this idea after searching literature and the internet for evidence that there is a need for his solution. He eventually found 4 pages of statements supporting his idea. **

[Time: 0.5/6.5] (4 points)

A He makes a fallacy

B He does not make a fallacy

19 *After choosing what idea it is going to be, he creates a plan of attack. In doing so, he looks at how classic business plans are put together. But because his idea is new, he finds those classic plans too rigid and decides not to completely fix everything yet. He works with mock-ups, checks at each step whether he can bear the risks of the next step, continuously checks whether the actions to be taken add more value than are sacrificed, tries things out and listens to feedback. After all, 'analysis paralyses. **

[Time: 0.5/6] (4 points)

A He makes a fallacy

B He does not make a fallacy

20 *The idea is taking more and more shape. For further elaboration, he looks for others who have qualities he does not have. For instance, he is looking for someone who can bring meticulousness and quality to the steering of development projects. Because a first meeting with someone he happened to meet at a trade fair disappointed him greatly, he decided to take a smarter approach: he started looking for self-employed people who had a very well-made Curriculum Vitae. **

[Time: 1/5] (4 points)

A He makes a fallacy

B He does not make a fallacy

21 *The freelancer he finally invites is a breath of fresh air compared to his first interview. Mainly because of this, he decides to hire this freelancer to help him in the project-based development of his idea.**

[Time: 0.5/4] (4 points)

A He makes a fallacy

B He does not make a fallacy

22 *He does a check that his idea complies with applicable laws and regulations. His approach takes these preconditions into account.**

[Time: 0.5/3.5] (4 points)

A He makes a fallacy

B He does not make a fallacy

23 *When testing his first real prototype, he looks for enthusiasts from the market he is targeting who can provide him with positive feedback. He gets energy and self-confidence from people like this.**

[Time: 0.5/3] (4 points)

A He makes a fallacy

B He does not make a fallacy

24 *He manages to realize at a 'mimimal viable product' and seeks investors. He eventually goes to a meeting where several wealthy 'Dragons' are looking for investment*

*opportunities and he gets a series of 5 interviews. The first three come to nothing, but this reinforces his idea that the chances at the last 2 are better: after all, the probability of a 'no' 5 times in a row is smaller than the probability of 1 or more of the 5 Dragons saying 'yes' (to be precise: the probability of 5 times no is 0.5^5, the probability of 1 or more times yes [1-0.5^5])..**

[Time: 1.5/2.5] (4 points)

A He makes a fallacy

B He does not make a fallacy

***25** During his pitch, the young entrepreneur was so confident that he made promises that are difficult to achieve and spoke in terms of '70% chance of winning' instead of '30% chance of losing'.**

[Time: 0.5/1] (4 points)

A He makes a fallacy

B He does not make a fallacy

***26** His startup is struggling after six months, which worries him, but he keeps going because he has already put so much effort and money into it. After all, the persevere wins..**

[Time: 0.5/0.5] (4 points)

A He makes a fallacy

B He does not make a fallacy

{After this test, the respondent immediately gets the obtained number of points. The right answers can be found in Appendix C. Probably the reader would like to take the test on

the fly. However, if the answers are given in a ready-readable row, the reader cannot prevent him/herself from reading the answer to the upcoming question beforehand. Therefore, the answers are giving in a matrix and can be found using the password "*althoughiknowyouwontcheat*".)

3.7.3 Measuring CT Conduct and Impact

Letting YPs report their own behavior is most likely not as accurate as letting others report their behavior (Chao and Lam, 2011). However, both ways of reporting may shed an accurate light on the behavior in reality (Vazire and Mehl, 2008). Therefore, the conduct of CT is both measured in one question by both the YP, and in detail in eights questions for the MO (see paragraph 3.2.2). For the (external) Impact measurement extra problems on top of inaccurate self-reporting occur. People probably overestimate the positive outcomes they achieve (Kruger and Dunning, 1999)(Kim *et al.*, 2016)(Dunning, Heath and Suls, 2004). It is also possible that YPs do not have a clear sight on the goal that their employers want to reach, as they tend to focus on their own growth (Zacher *et al.*, 2009). Therefore, it might be problematic to ask the YP what should, in the eyes of the organization, be deemed as "important". In the definition of Impact in this report however, the notion what is important to the employer is necessary to define what Impact in this situation means. Therefore, the measurement of the Impact outside the organization the YP makes, is solely asked to their manager.

The logic behind the measurement is discussed in paragraph 3.2.2.

3.7.4 Stimulating CT

A training for YP's was planned, where the CT Ability, Opportunity and Conduct would be stimulated. The idea was to perform a before- and after measurement on Impact, to see if CT caused Impact. In a two-moths period 4 sessions would be held. For every session the participating YPs would have to read some theory and make a homework assignment. This homework assignment would be directly applicable to the working environment of the YP, thus "forcing" the YP to conduct CT behavior. This assignment would be discussed at the next training.

Based on CT training literature (Catchings, 2015)(Helsdingen *et al.*, 2010)(*Red Teaming Handbook*, 2021)(Hoffman, 2017)(Passmore and Rowson, 2019)(Vlerick, 2022)(Sandra P.A. Robinson and Verna Knight, 2019)(Senge, 2014)(Beck and Cowan, 2014) and the experience in training, consultancy and didactical material the researcher collected during his career, a comprehensive program would be composed. This would encompass theory of CT and Thinking Biases and tools for analyzing CT challenges, challenging thinking models of others on individual and team level (Metamodel, Socratic techniques, Facilitation Techniques, advising visually), Systems thinking (Senge's archetypes) and scenario analysis, as well as connect and communicate with others using human drives (Spiral dynamics).

The idea was to let the participants evaluate the usefulness of the tools and techniques presented in the training "t-1", and "vote" for the topics to come, thus making it a practical research program.

Also, participants would be stimulated to find a "Mentor" in their working environment, who would be no part of the program. Unlike "normal" programs where the Mentor coaches the YP, in this trajectory the YP would also be invited to challenge the Mental Model of the Mentor, not by "telling" alternative truths, but by questioning

his/her Mental Model in a Socratic manner. This way, in direct connection with persons of influence, the possibility of a YP making Impact would hopefully be maximized.

However, in absence of enough participants, this training program is not developed further.

3.8 Population, Sample and Participant Selection

The population this research project aimed on, are Young Professionals that completed a HEI study recently, and their managers. As we will see in paragraph 3.9, the focus as described in the current paragraph 3.8 had to be expanded, due to the low number of YPs that wanted to take part in this research project.

- Although it is a kind of "convenience-sampling" that should be avoided, for practical reasons samples were initially taken from the student-population of the RBS. This would increase the relevance of the study-outcome for the RBS, but at the same time decreases the validity of generalization of the results.
- Not every graduate will work immediately after graduation at a university of applied sciences in their first job. About 23% will continue their study at a science university. About 4% will not have a job within one-and-a-half years. However, about 87% of the Young Professionals that do enter the labor market start immediately after (or sometimes even before) graduation. And 93% will have found their first job within three months (*Kansen op de arbeidsmarkt voor afgestudeerde hbo'ers*, no date). RBS-students are likely to switch to another job within 2 years (Statistiek, 2023). Job switchers are likely to take some months to orient. It isn't fair to expect YPs to make Impact within a week. It is not fruitful to start with YPs that are likely to switch jobs during the research project either. So, to take some margin, the "overdue date" starts at two years minus three months

orientation phase minus the half year period the research project takes. So, the Survey would not be sent earlier than three to six months after the graduation ceremony. But also, not later than thirteen months after that ceremony in order to "miss" the period that they are searching for a new job.

- The focus of the study are young HEI-graduated professionals. The Rotterdam Business school has several studies, some of them (5 in total) are aimed at working professionals and are given parttime. The part-time studies were therefore excluded from the sample. Samples would be drawn from the alumni-pools from fulltime studies at bachelor level (14) and Master level (2).
- There is evidence that a certification program makes a training more appealing (Rodrigues, 2019)(Andresen, 2019). In order to boost participation, the researcher made sure a "official certificate" could be obtained by YPs that would participate. After a study what trademark would be appealing and would not conflict with existing trademarks, the "Certified Tenth - C10th" trademark was chosen. C10th stems from the idea of the "tenth man doctrine" and is a nod to CT without calling it CT (Critical Tenth). The trademark was registered at the EUIPO, making it "official", getting the possibility to allow others to use it as a title, and prohibiting misuse (Figure 7).

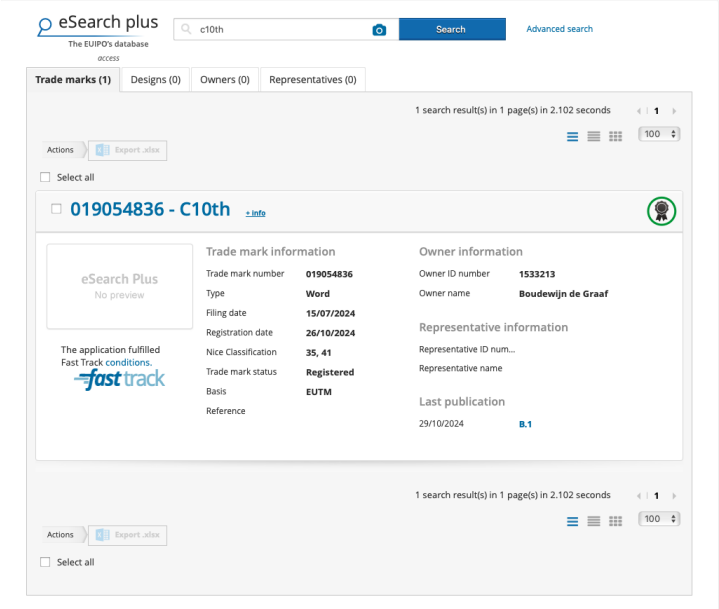


Figure 7:
Registrated trademark C10th

- A logo was designed by a specialized logo designer. Also, two domains were claimed (https://www.c10th.com and https://www.c10th.org). A website was realized by the researcher, using Dall-E for generating the pictures and Carrd.com for the basic design (Figure 8).

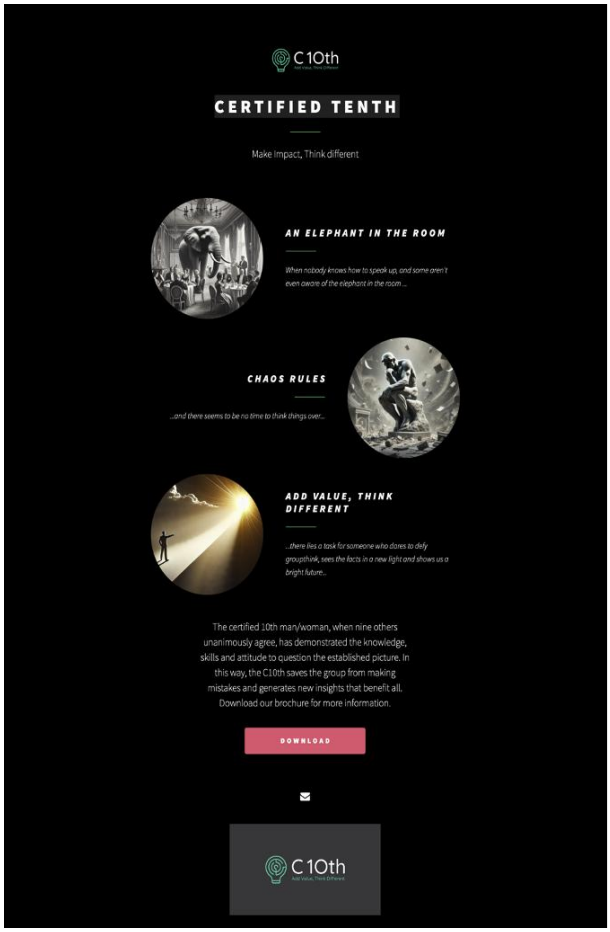


Figure 8:
Logo and Website https://www.C10th.org

- YPs were notified of the research at their graduation ceremony with a leaflet (Figure 9) saying (Translation by Deepl.com and manually corrected): “As a newly graduated Young Professional, you want to make Impact. You think at HBO level. But what does this way of thinking bring you and your organisation exactly? We are looking for Young Professionals who are interested in participating in a study to find out. This research will start sometime around December and consists of a questionnaire to be completed by you and your manager. Among the respondents, 15 places will be chosen for a free ‘Red

Teaming' course, where upon successful completion you will be awarded the protected title of 'Certified Tenth' (C10th) after your name. Would you like to be approached in a few months to participate in this survey, please leave then leave your Email address:"

Figure 9:
Leaflet

- The YPs were notified again for the second time after 3 months by email (when was to be expected that they found their first Job).

3.9 Data Collection

The actual data collection didn't go as foreseen and forced the researcher to broaden the search and the population. The original idea was to aim on freshly graduated HEI-alumni from the RBS. However, due to the low response, the final population was HEI-alumni that graduated 8 years or less. The next steps were undertaken:

- The researcher made sure that leaflets and posters were available at all of the RBS graduation ceremonies. Graduates could leave their email in a MS-Forms form and that way show their interest in participating.
- There was also a short live introduction held at the graduation ceremony of the Business & It Management study, part of the RBS
- After it became clear the number of potential participants would be minimal, the researcher broadened the population and aimed at all Dutch HEI graduates, that graduated 8 years or less ago.
- The alumni groups of the RBS, if they existed, were contacted via the managers of the study areas.
- LinkedIn was used to reach candidates. A message was spread to the contact of the Researcher (500+ contacts), asking to spread the message further. This message got 1245 impressions, reaching 574 members. 82% of them lived in The Randstad, 36% were at senior level ((*LinkedIn Analytics*, no date).
- The researcher contacted 63 YPs -former students from his own network. All students were contacted by name and a written message especially for each contact.
- Also, a trial version of LinkedIn was used to contact possible participants by name that were no contact of the researcher yet.
- The account manager of the official RBS channel copied the message, thus reaching the business partners of the Rotterdam Business School.

At the end 31 unique YPs responded via the Survey, and 13 managers. However, it was not possible to match 2 managers to an YP.

In the messages aimed at the YPs the timeframe for the C10th training was already mentioned. Once it was clear to the researcher that 19 YPs wanted to attend the C10th training part, but only 4 of them could be matched to a manager, the idea of a before- and after measurement was abandoned. Offering a training for 4 YPs, even if they all would attend, makes the basis for conclusions too small. Therefore all 19 YPs that showed interest were contacted with concrete dates for the training, starting 6 weeks after this message.

The researcher hoped to find a way to do a single "delta measurement" after the training and get an idea of how valid it was when it could be compared to a before and after measurement from the few YPs that did have a manager that filled in the before-measurement. Knowing that it would be not as ideal as a before- and after measurement for all if the YPs.

Possible due to the "planning fallacy" bias, only 2 YPs signed in for the C10th training. To the opinion of the researcher this number of YPs would not lead to valid and reliable results. The researcher contacted some YPs and got the feedback that the course conflicted with other activities like sports, and also seemed quite a lot of work in a relative short period. The timeframe for this research however made it impossible to find another workable solution. Therefore, this this planned part of the research project was abandoned.

3.10 Data Collection Procedure

Data was collected via MS-Forms. Every respondent agreed to the Consent Form, that was made available online via https://edu.nl/fypvv. In order to gain more

trustworthiness, the "real" proton url pointing at the Consent Form was shortened with a specialized url-shortener for research purposes (*edu.nl - Dé URL-shortener voor onderwijs en onderzoek met respect voor privacy.*, no date). In this form the participation, rights and procedures were communicated in both English and Dutch. This consent form gives a detailed description of the procedures followed, and rights of the respondent. It follows the lay-out prescribed by the Swiss School of Business Management, Geneva. The English text of the consent form is added as Appendix D.

Participants were led to the overarching survey. The first question was to fill in the email address of the YP, in order to be able to match the answers. The next question was if the respondent was a Young Professional, of his/her Manager or Senior (other). Depending on the answer the respondent was led to the YP section, or the MO section.

The MO section was quite straightforward and at the end the MO was asked if he/she would like to receive the results of this research and was thanked for participating.

The YP section was for the first part quite straightforward also, but at the end the YP was asked to take the CT ability test in a separate form (link included in the first form, that opened a new window). After completing the CT test, the YP could fill in the obtained points in the overarching survey. Then the YP was asked if he/she already noticed his/her MO, and was given a suggestion for an email to his/her MO. At the end of the survey the YP was asked if he/she would like to receive the results of this research and was thanked for participating.

Hoping to get more "traction" the researcher thought it would be a nice idea to put a "gaming" element in the process. The YP and MO could do a "battle" to find out if the YP scored higher in CT Ability then the MO. None of the MOs used this possibility.

In some cases, the YP filled in his/her personal email address and the MO his/her work email address, making it impossible to match the two instances. The researcher sent an email rappel twice to the addresses that could not be matched, and at the end only two MO reactions could not be matched to a YP.

3.11 Data Analysis tools

The next tools were use for data analyses:

- For the survey and downloading the results in an Excel-readable format: MS-Forms.
- For cleaning the data and converting it to a format that could be used for statistical analyses: a by the researcher hand-build "translator" in MS-Excel.
- For the "raw" statistical analysis: Real Statistics (*Home Page (Welcome) | Real Statistics Using Excel*, no date) and MS-Excel.
- For double-checking the Real Statistics outcomes and generating Harvard-style tables and Figures: Jamovi (*jamovi - open statistical software for the desktop and cloud*, no date).

3.12 Research Design Limitations

The following limitations were recognized beforehand:

- The used thinking model was plausible but not tested yet. The same goes for the definitions. If the data indicates clearly that the thinking model is flawed, the research question may not be answerable using the obtained data.
- The number of respondents is probably low, making the results unreliable.

- The data collection was partly conducted with “snowballing”, what probably makes the sampling biased.
- Without the training-part, it will not be possible to test for “causation” between CT and Impact
- Because of the limitations, like the time that conducting a survey may take at maximum, it is impossible to address all the elements in the original thinking model. Therefore, a simplification was put in place. Effects on the “Mental Model” of the group and organization are “measured” and defined as “Impact”. Effects on the Mental Model of the YP is not measured. Even when there would be enough participants, the original thinking model shall not be tested in full.

3.13 Conclusion

Using the theory and thinking model from Chapter II, this chapter described what in this project is meant by CT and Impact. To find an answer on the main research question if YPs with CT can make Impact a survey is conducted. Due to the lack of participants and the limited timeframe, the foreseen training to find out if there is not only a correlation, but also a causation relation between the two concepts, had to be abandoned. In the next chapter the answers of the 44 respondents will be analyzed in detail. In Chapter V the analysis will be interpreted.

CHAPTER IV:

RESULTS

Using the theory and thinking model from Chapter II, and the operationalization and actions as described in Chapter III, this chapter describes the results of the conducted Survey. First some descriptive statistics are given. Then the usability of the Simplified Thinking Model is checked. After that a diversity of correlations are analyzed. The chapter concludes with a summary of the findings.

4.1 Descriptive Statistics

4.1.1 Number of Respondents

The number of respondents on the survey were build up as follows:

- The survey had 44 respondents in total, including 13 MO and 31 YPs
- Out of the 13 MO respondents, 11 could be matched with an YP, leaving 2 MOs unmatched
- Out of the 31 YP respondents, 11 could be matched, leaving 20 YPs unmatched.

4.1.2 Number of Daily Colleagues

Nearly half of the YP's worked on a daily basis together with 1 to 5 colleagues (YPTeamSize coded as "1"), one-third in a team size 6 up to 10 (coded as "2"), and nearly one-fifth with 11 to 20 colleagues (coded as "3"). No one worked in bigger teams (Table 1).

Table 1:
Teamsize of the YPs Daily Work Environment

Frequencies of YPTeamSize

YPTeamSize	Counts	% of Total	Cumulative %
1	14	45.2%	45.2%
2	11	35.5%	80.6%
3	6	19.4%	100.0%

4.1.3 Levels of Education

The highest level of education of the YP's that responded, was quite evenly distributed on a University of Applied Sciences (labeled as "HBO", 45.2%) and Research University (Labeled as "WO", 54.8%). The education levels of the matched MOs were more or less distributed the same (Table 2 and Table3).

Table 2:
Number of YP Respondents split by Education Level

Frequencies of YPEduLevel

YPEduLevel	Counts	% of Total	Cumulative %
HBO	14	45.2%	45.2%
WO	17	54.8%	100.0%

Table 3:
Number of MO Respondents split by Education Level

Frequencies of MOEduLevel

MOEduLevel	Counts	% of Total	Cumulative %
Anders	1	9.1%	9.1%
HBO	5	45.5%	54.5%
WO	5	45.5%	100.0%

4.1.4 Months Working

The period YPs worked after their graduation has a mean of 32.1 months (nearly 2.5 years), with quite a great standard deviation of 25.8 months (Table 4). The “hilltops and valleys” in the graph (Figure 10) are not totally unexpected. Given that most of the Dutch studies know a 12 month cycle, that most of the YPs will have found a job after 3 months of their graduation and that the measurement took place halfway a cycle, some fading out cyclic effects could be expected.

Table 4:
Months Working of YPs after Graduation

Descriptives

	YPJobMonths
N	31
Missing	0
Mean	32.1
Median	35
Standard deviation	25.8
Minimum	0
Maximum	88

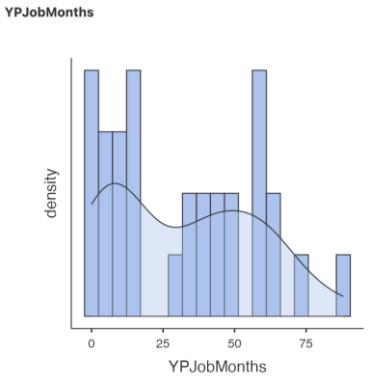


Figure 10:
Months YPs working after Graduation

4.1.5 Benefits of the HEI education

The profits that respondents could choose were “Diploma” (Giving access to a job at the desired level), “Knowledge” (knowledge of the field of expertise of the study), “Think” (the HEI-way of thinking, a CT description used by the researcher when trying to avoid the negative connotated word “Critical” in the survey) and “Network” (acquaintance to others). The “Other” benefit field was used by no one. The overall picture of all 44 respondents (Table 5) was that “Diploma” was the most important (77.27%), Then came “Think” (70.45%), and at a distance “Knowledge” (59.09%) and “Network” (31.82%). This order of importance was independent of the respondent being MOs or YPs. A difference is seen when we split up the group in type of education. Using a one-tailed two-sample proportion Z-test, we can conclude that with 12 successes out of 23 WO (Research university) respondents scoring on “Knowledge” as a benefit, against 16 out of 20 HBO (Applied sciences) respondents doing the same, the HBO respondents score significantly higher on this benefit (p = .028). It is likely that Alumni from research universities, looking at the type of organization of the respondents, did not end up in a research position, do not choose “Knowledge” as very beneficial.

Table 5:
% scored Benefits of HEI education

	#Respondents	Dipl	Think	Knowledge	Network
ALL	44	77,27	70,45	59,09	31,82
MO	13	61,54	61,54	61,54	38,46
YP	31	93,55	80,65	64,52	38,71
HBO	20	85,00	75,00	80,00	40,00
WO	23	86,96	78,26	52,17	39,13

Therefore, it may be concluded that, according to the respondents, CT is considered as a valuable competence that was learned at their study.

4.1.6 Impact on the Environment

The YPs felt, with an overall mean of 4.1 and a standard deviation of .7 on a 5 points scale, that they mostly had a significant Impact on their working environment. No one felt they had no Impact (Table 6 and Table 7).

Table 6:
Overall Impactscore YP, self-report

Descriptives

	YPImpactEnvir
N	31
Mean	4.10
Median	4
Standard deviation	0.700
Minimum	3
Maximum	5

Table 7:
Overal Impactscore YP, self-score, Distribution

Frequencies of YPImpactEnvir

YPImpactEnvir	Counts	% of Total	Cumulative %
3	6	19.4%	19.4%
4	16	51.6%	71.0%
5	9	29.0%	100.0%

The MOs, asked about the Impact “their” YPs made on their environment, with a mean of 4.0 and a deviation of .9, did mostly agree with this (Table 8 and Table 9).

Table 8:
Overall Impactscore YP, MO-report

Descriptives

	MOImpactEnvirYP
N	11
Mean	4.00
Median	4
Standard deviation	0.894
Minimum	3
Maximum	5

Table 9:
Overall Impactscore YP, MO-report, Distribution

Frequencies of MOImpactEnvirYP

MOImpactEnvirYP	Counts	% of Total	Cumulative %
3	4	36.4 %	36.4 %
4	3	27.3 %	63.6 %
5	4	36.4 %	100.0 %

A two-tailed Kendall's tau correlation analysis was conducted to investigate the relationship between YPImpactEnvir and MOImpactEnvirYP for the 11 paired MOs and YPs. A strong positive statistically significant relationship has been found, $\tau = 0.707$, $p = .018$, n=11. This suggests that the scores given by the YP and the MO for their Impact strongly correlate but are not exactly the same.

To find out if the difference between the two values is significant, a two-tailed Wilcoxon Signed Rank Test was conducted to test if the YPImpactEnvir and MOImpactEnvirYP for the paired MOs and YPs differed. The results show a non-significant result, with $W = 9$, $p = .25$, $n = 11$, where the median of MOImpactEnvirYP was 4 and the median of MOImpactTotal was 3. It may be concluded that the YP and MO make the same estimates about the Impact the YP makes.

Does this mean that this MOImpactTotal score can be used as a strong predictor for the calculated Impact score as defined in this research? Because only the MOs score

the input for the calculated Impact (MOImpactTotal), the correlation between MOImpactTotal and MOImpactEnvirYP is analyzed.

A two-tailed Kendall's tau correlation analysis was conducted to investigate the relationship between MOImpactTotal and MOImpactEnvirYP for the 11 paired MOs and YPs. A weak positive statistically unsignificant relationship has been found, $\tau = 0.271$, $p = .309$, $n = 11$. This suggests that the MOImpactEnvirYP question cannot be used as a "single question" after the training. Also, it is possible that the idea of what Impact encompasses is, to the MO, not "the same enough" as how it is defined in this research.

4.1.7 CT Opportunity

YPs scored with medians of 4 on the feeling of the inclusion in the team and the freedom to reinvigorate (Table 10). The MOs scored both variables with a median of 5, asked about these scores for the YPs. The MO's scored slightly higher on the "mean" of freedom to be critical on the status quo (both medians were 4).

Table 10:
CT Opportunity related scores, all YPs and MOs

Descriptives

	YPBelonging	YPCTAllowed	YPCTFreedom	MOBelongingYP	MOCTAllowedYP	MOCTFreedomYP
N	31	31	31	11	11	11
Missing	0	0	0	20	20	20
Mean	4.13	4.06	4.26	4.36	3.91	4.55
Median	4	4	4	5	4	5
Standard deviation	0.718	0.854	0.729	0.809	0.944	0.522
Minimum	2	2	2	3	2	4
Maximum	5	5	5	5	5	5

Looking at solely the YPs that had a matching MO (Table 11), the mean "Freedom" is scored higher by the MOs than by the YPs:

Table 11:
CT Opportunity related scores, Matched YPs and MOs only

Descriptives

	YPBelonging	YPCTAllowed	YPCTFreedom	MOBelongingYP	MOCTAllowedYP	MOCTFreedomYP
N	11	11	11	11	11	11
Missing	0	0	0	0	0	0
Mean	4.18	3.82	4.36	4.36	3.91	4.55
Median	4	4	4	5	4	5
Standard deviation	0.405	0.603	0.505	0.809	0.944	0.522
Minimum	4	3	4	3	2	4
Maximum	5	5	5	5	5	5

This means that the MO's probably feel the opportunity to conduct CT for the YPs is higher, than the YPs feel for themselves.

A statistical test was performed. Considering the times MOs or YPs score individually higher on Belonging, CTAllowed and CTFreedom than their "peer", 33 cases can be compared (11 paired MOs/YPs, times 3 items). In this case the MO scores 14 times higher than the YP, and the YP scores 6 times higher than the MO. Both give 13 times the same score (making a total of 14+6+13=33 comparisons). Using a one-tailed two-sample proportion Z-test, we can conclude that with 14 successes out of 33 possibilities for the MOs, against 6 successes out of 33 possibilities for the YPs, the MOs score significantly higher than YPs (p = .016) on the opportunity the YP has to conduct CT thinking.

4.1.8 Reflection Mode

YPs see their mode of reflection with a mean of 3.55, at a five-point scale, slightly higher as the MOs scored their reflection mode with a mean of 3.36, see Table 12.

Table 12:
Reported Reflection Mode by YPs and MOs

Descriptives

	YPReflectionMode	MOReflectionModeYP
N	11	11
Missing	0	0
Mean	3.55	3.36
Median	4	4
Standard deviation	0.820	1.21
Minimum	2	2
Maximum	5	5

To find out if the difference between the YP and MO score is significant, a two-tailed Wilcoxon Signed Rank Test was conducted to test if the YPReflectionMode and MOReflectionModeYP for the paired MOs and YPs differed. The results show a non-significant result, with $W = 24$, $p = .770$, n = 11, where the median of YPReflecttionMode was 4 and the median of MOReflectionModeYP was 4. This means that the YP and MO make the same estimates about the reflection mode of the YP.

Another statistical approach leads to comparable results. Considering the scores for each of the 11 YP/MO pairs, the MO scores 4 times higher than the YP, the YP scores 6 times higher than the MO, and both have 1 time the same score (making 6+4+1=11 comparisons in total). Using a one-tailed two-sample proportion Z-test, it may be concluded that with 6 successes out of 11 possibilities for the MOs, against 4 successes out of 11 possibilities for the YPs, the hypothesis that the differences is not rejected ($p = .19$). The differences may be purely coincidental.

4.1.9 Miscellaneous MO numbers

The data led to the next observations:

- All but 1 MO was the manager of the YP.
- 7 of the MO/YP pairs were working in a MOISIC K62 organization ("Computer programming, consultancy and related activities"), 2 in a MOISIC L organization ("Financial and insurance activities"), 1 in a MOISIC P organization ("Public administration and defense; compulsory social security") and 1 in a MOISIC R organization ("Human health and social work activities")
- All but 1 MO/YP pairs worked in a Large SME (250 or more employees, more than 50 million turnover), 1 pair worked in a small SME (less than 50 employee, maximum turnover of 10 million)

It may be concluded that most of the YPs were scored by their manager, and that overall, the YPs worked in a Large SME in the Branche "Computer programming, consultancy and related activities". It will, due to the limited variance, not be very fruitful to try to break up the analysis to Branche or organization size.

4.2 Usability of the Thinking Model for the Calculated Variables

In chapter 3.4.1 predictions were made based on the Simplified Thinking model. Before using the calculated variables, it is wise to get an idea about the "power" of this model. If the measurements contradict the predictions, we may conclude that the Simplified Thinking model, and the way the calculations are set up are biased.

4.2.1 Usability of the Thinking Model for the Calculated Variables

In Chapter 3.4.1 it was argued that the next predictions could be used to falsify the Thinking Model as developed in Chapter II.

1. In the Mental Model, it is more difficult to make Impact on the "Self" than Impact on the feeling and perceiving. The number of reported scores on both columns MOImpactOrgFeeling and MOImpactOrgPerceiving should individually exceed the number of Impact scores on MOImpactOrgSelf. We can see (Table 13) that there is only 1 time the MO scored the column MOImpactOrgSelf (value = "WAAR"), and in that scored row the MOImpactFeeling was also scored "WAAR" (="TRUE"). In total the Feeling and Perceiving were 5 times scored. The prediction stemming from the Mental Model is therefore not falsified

Table 13:
Impactscores MO on the Mental Model dimensions

MOImpactOrFeeling	MOImpactOrgPerceivi	MOImpactOrgSelf
WAAR	ONWAAR	ONWAAR
WAAR	ONWAAR	ONWAAR
WAAR	WAAR	ONWAAR
ONWAAR	ONWAAR	ONWAAR
ONWAAR	WAAR	ONWAAR
ONWAAR	ONWAAR	ONWAAR
ONWAAR	WAAR	ONWAAR
ONWAAR	WAAR	ONWAAR
WAAR	ONWAAR	WAAR
WAAR	WAAR	ONWAAR
5	**5**	**1**

2. It is more difficult to make Impact externally than internally. So, the number of reported scores on the Mental Model as a whole MOImpactOrgTotal), should exceed the number of reported scores on the

External Impact (MoImpactEnvirTotal). N.B.: notice that both variables consist of 3 sub variables, which makes “normalization” unnecessary. The number of times Impact scored more “WAAR” then Org is 1 against 10. Also the total number of “WAAR” for the SELF column scored is 10, against 7 times for the Envir column (Table 14). So, the prediction stemming from the Mental Model is therefore not falsified

Table 14:
Impactscores MO on both Mental Model and Environment dimensions

MOImpactOrFeeling	MOImpactOrgPerceivi	MOImpactOrgSelf	#	MOImpactEnvirH0	MOImpactEnvirH1	MOImpactEnvirH2	#	Org>=Envir?
WAAR	ONWAAR	ONWAAR	1	WAAR	ONWAAR	ONWAAR	1	WAAR
WAAR	ONWAAR	ONWAAR	1	ONWAAR	ONWAAR	ONWAAR	0	WAAR
WAAR	WAAR	ONWAAR	2	WAAR	ONWAAR	ONWAAR	1	WAAR
ONWAAR	ONWAAR	ONWAAR	0	ONWAAR	ONWAAR	ONWAAR	0	WAAR
ONWAAR	WAAR	ONWAAR	1	ONWAAR	WAAR	ONWAAR	1	WAAR
ONWAAR	ONWAAR	ONWAAR	0	ONWAAR	ONWAAR	ONWAAR	0	WAAR
ONWAAR	WAAR	ONWAAR	1	WAAR	ONWAAR	ONWAAR	1	WAAR
ONWAAR	WAAR	ONWAAR	1	WAAR	WAAR	ONWAAR	2	ONWAAR
WAAR	ONWAAR	WAAR	2	WAAR	ONWAAR	ONWAAR	1	WAAR
WAAR	WAAR	ONWAAR	2	WAAR	ONWAAR	ONWAAR	1	WAAR
			10				7	

3. In the external Impact, the “lower levels” are easier to influence than the “higher levels”. So, the number of reported scores of Impact on a lower level should exceed the number of reported scores on the higher levels (#scores on MoImpactEnvirHa > #scores on MoImpactEnvirHb, where a < b). The number of scored MOImpactEnvirH0 is 6, exceeding the number of scores of MOImpactH1 which is 2. MOImpactEncirH2 is scored 0 times (Table 15). So, the prediction stemming from the Mental Model is therefore not falsified

Table 15:
Impactscores MO on the Environment dimensions

MOImpactEnvirH0	MOImpactEnvirH1	MOImpactEnvirH2
WAAR	ONWAAR	ONWAAR
ONWAAR	ONWAAR	ONWAAR
WAAR	ONWAAR	ONWAAR
ONWAAR	ONWAAR	ONWAAR
ONWAAR	WAAR	ONWAAR
ONWAAR	ONWAAR	ONWAAR
WAAR	ONWAAR	ONWAAR
WAAR	WAAR	ONWAAR
WAAR	ONWAAR	ONWAAR
WAAR	ONWAAR	ONWAAR
6	**2**	**0**

4. According to the thinking model, Impact on the Mental Model proceeds Impact extern. So, it is possible to make Impact on the Mental Model without making Impact extern, but it is mandatory to make Impact on the Mental Model in order to make Impact in the outside world. So, there should be no cases where MoImpactEnvirTotal is scored, and MoImpactOrgTotal isn't. But it is possible that there are cases where MoImpactOrgTotal is scored, but MoImpactEnvirTotal isn't. Using the table (Table 14) as given in prediction 2, we can see that there is 1 case where Impact on the Organization level is reported, and not at Environment level, and that there is no case where Impact is reported at Environment level but not at Organization level. So, the prediction stemming from the Mental Model is therefore not falsified.

It can be concluded that all predictions stemming from the thinking model hold, and therefore we don't reject the Thinking Model.

A question that remains is whether the proposed weights are reasonable compared to the reality. Tho check this, an expectancy table is made using the assumption that it takes more effort to make big Impact, than to make a small Impact. When the formula

would be “EFFECT = 2A+B”, where the weight of A is 2, the expectation would be that A is 2 times more difficult to reach than B (as 1 A equals 2 B’s to reach the same EFFECT result). So, in this case, using the assumption, we would expect 2 B’s scored against every 1 A. Using this logic, an expectancy table is made for the frequencies, where the measured results can be tested against (Table 16). A chi-square Goodness of Fit test was conducted to determine whether the observed frequency counts differ significantly from the expected values. The results were statistically unsignificant, with $\chi 2(5) = 6.021$, $p = .304$, $n = 60$. We may conclude that the deviance in the measured values fall within the range of coincidence. This means that the “weights” as proposed can be used in the rest of this research project.

Table 16:
Expectancy table Weight-based Frequencies and Measured Frequencies

		MOImpactOrgTotal			MoImpactEnvirTotal		
	EXPECTED						
	MOOrgPerc	MOOrgFeeling	MOOrgSelf	EnvirH1	EnvirH2	EnH3	Total
#Scored	5	5	2	3	2	1	18
#NotScored	5	5	8	7	8	9	42
Total	10	10	10	10	10	10	60
	MEASURED						
	MOOrgPerc	MOOrgFeeling	MOOrgSelf	EnvirH1	EnvirH2	EnH3	Total
#Scored	5	5	1	6	2	0	19
#NotScored	5	5	9	4	8	10	41
Total	10	10	10	10	10	10	60

4.2.2 Usability of the Ability Test

In Chapter 3.7.2 the “making of” the CT Ability test is described. In this paragraph a final check is conducted to find out if the test held up in “real life”. The data led to the next observations:

- 33 respondents took the test (1 respondent took the test twice)

- The mean time the test took everyone was 14.3 minutes, $SD = 4.02$, $n = 33$, with a more or less normal distribution (Shapiro-Wilk $W = .940$, $p = .069$). The median was 15.1 , making the histogram a bit left-skewed.
- The Cronbachs' alpha $\alpha = .644$ and could be brought up to .722 when questions with a negative item-rest correlation would be omitted. However, this score is "good enough" as it is, and it is explainable why the questions that suppress the alpha score a bit, were scored low.
- Questions 3, 5 and 12 are the more difficult logical reasoning questions, which could indeed be seen in the score.
- For question number 15, recognizing the Survivors bias, was the low score a bit expected, as YPs may be used to following "Influencers".
- For question number 25 the low score for recognizing the bias was also a bit expected. YPs may have learned that manipulation / making things look better than they are, is part of "advertising".

It can be concluded that the Ability test performed well (enough), and gives a workable impression of the respondents' ability to conduct CT.

4.3 The Relationship between CT and Impact

4.3.1 CT Ability of HBO and WO compared

To find out if Research University (WO) YPs and Applied Science University YPs score the same on the CT Ability tests, several tests were performed.

The average CT Ability score (YPChallengeScoreReported) of the YPs is 71,17. The type of education is described in YPEduLevel. Of the 16 YP's with a Research University (WO) background, 13 scored higher than the average. Of the 13 Applied Sciences (HBO) graduated, 6 scored higher than the average.

Using a one-tailed two-sample proportion Z-test, and the variables YPEduLevel and YPChallengeScoreReported, we reject that these scores are based on coincidence (p = .023). We may assume that Research University YPs perform higher in the CT Ability test and therefore have a greater CT Ability than the YPs with an Applied Sciences background.

Also, a two-tailed independent t-test was conducted on the same data to find out if there is a significant difference between the WO and HBO YPs on YPChallengeScoreReported. The results showed a significant difference, $t(27) = 3.43$, p = .002, d = 1.284. The average score for the WO was 78.25, SD = 8,13, while the average score for the HBO was 62.46, SD = 16.04. This test too shows that we may assume that Research University YPs perform higher in the CT Ability test and therefore have a greater CT Ability than the YPs with an Applied Sciences background.

4.3.2 CT Conducting of HBO and WO compared

We saw in the former paragraph that Research University YP's have a higher CT Ability than Research University (HBO) YPs. But do they also score higher on conducting CT?

The average score that the 11 MOs gave YPs for CT behavior (MOBehaveYPTotal) is 82. From the 7 WO YPs, 4 scored higher than the average. From the 4 HBO YPs, 2 scored higher. Using a one-tailed two-sample proportion Z-test, we do

not reject that this difference is within the range of coincidence (p = .409). We may therefore NOT assume that Research University YPs conduct more CT than YPs with an Applied Sciences background.

Also, a two-tailed independent t-test was conducted on the same data to find out if there is a significant difference between the WO and HBO YPs on MOBehaveYPTotal. The results showed a not significant difference, $t(9) = 0.15$, p = .883, d = 0.095. The average score for the WO was 81.57, SD = 10.03, while the average score for the HBO was 82.75, SD = 16.15. This test shows too that there is no difference in the conduct of CT between WO and HBO YPs.

4.3.3 Conducting CT and Making Impact: Correlation

Does Conducting CT (MOBehaveYPTotal) correlate with making Impact (MOImpactTotal)? A Pearson correlation test was conducted. The results showed a positive significant correlation between the two scores with $r(10) = 0.512$, p = .044. This means that a high level of conducting CT indeed correlates with a high score for making Impact (Figure 11).

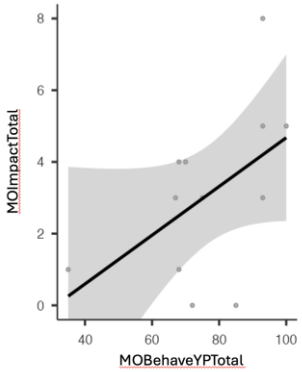


Figure 11:
Correlation between Conducting CT and Impact

4.3.4 Conducting CT and Making Impact: Causation

Due to the missing of the training part in this research, and therefore the lack of before- and after measurements of Impact after stimulating YPs to conduct CT behavior, a pure causation cannot be proven with the data from the survey.

But also, the shift in definition of what an YP is had consequences. In the original setting of this research YPs with a maximum of 2 years' work experience would enroll. In this setting it was not likely that Impact should have influence on CT Behavior. However, this definition was abandoned. With the stretching of the definition of YPs up to 8 years' work experience, it is possible that there is a kind of "success to the successful" loop introduced. Conducting CT may enhance Impact (through changing Mental Models), but also Impact may enhance Critical Thinking (e.g. by boosting the self-confidence of YPs, and/or the trust they gain from the colleagues).

It may be nice to find indications whether there is such a loop visible, and if so, what would be the natural "leverage" to speed up the circle. So, is it possible to say more about relationship between conducting CT (MOBehaveYPTotal) and the Impact the YP makes (MOImpactTotal), and vice versa[6].

First, does Impact and/or Conducting CT possibly improve over time (YPJobMonths). (N.B: because of 1 missing YPJobMonths-score, we work with 10 pairs of MO/YP pairs). Pearson correlation tests were conducted. The results showed a significant positive correlation between YPJobMonths and MOBehaveTotal with $r(8) = 0.68$, $p = .015$, and a significant positive correlation between YPJobMonths and MOImpactTotal with $r(8) = 0.64$, $p = .023$ also. This means that we may conclude that

[6] This is not a "standard" way of testing, and the researcher did not find literature about the logic that follows. The outcomes should therefore be handled with care.

the longer the YP is in a working environment, both the reported CT behavior and the reported Impact improve (Figure 12).

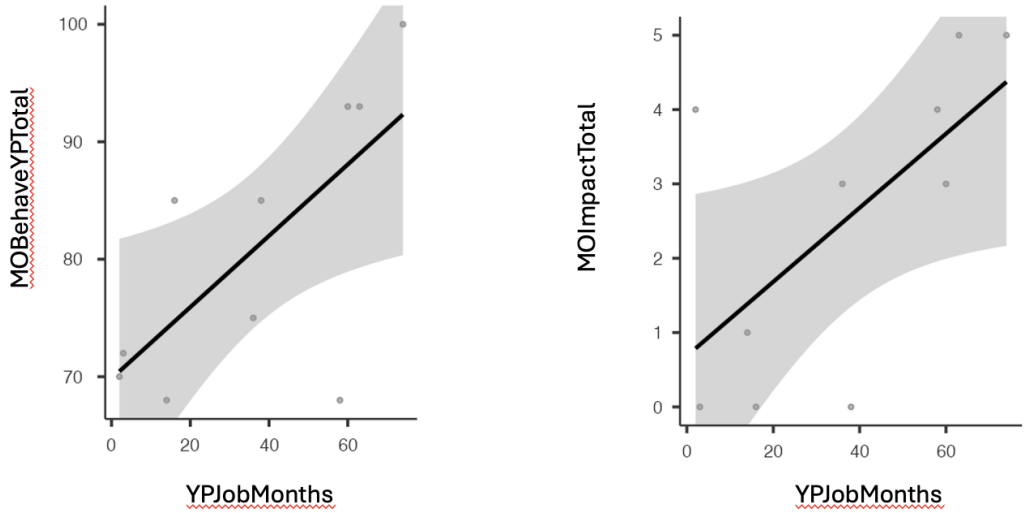


Figure 12:
Growth CT Conduct and Impact over time

Does the CT Ability also improve over time? A Pearson correlation test was conducted. The results showed no significant correlation whatsoever between YPJobMonths and YPChallengeScoreReported with $r(26) = 0.02$, $p = .918$. So, we may conclude that there is no CT Ability Improvement over time.

Within the logic of the Thinking Model, the "opportunity" may have grown in time. The correlation between YPJobMonths and the sum of YPBelonging, YPCTAllowed and YPFreedom ("Opportunity") was sought, and a Pearson correlation test was conducted. The results showed no significant correlation whatsoever between the two variables with $r(28) = 0.08$, $p = .346$. So, the (feeling of) opportunity does not grow over time.

Outside the Thinking Model other processes may have occurred, like choosing another job when an YP does not make Impact, gaining a higher hierarchical position and therefore have more influence, earning the trust of colleagues and therefore making it easier to influence their Mental Model, a regressive bias, physical maturing of the brain, et cetera). There are studies on what motivates YPs to choose their first job and accelerates the growth of Young Professionals. Personal traits and mentorship are amongst them (Skrobach *et al.*, 2020) (Franko and Myttseva, 2022)(Kolesnikova, 2019). But it is unlikely that "personal traits" or "mentorship" grow over time.

However, whatever the underlying cause may be, within the model we can ask whether the "mode" of conducting CT changes over time. Could it be that in the early months the YP is more self-oriented, and in the later months he is "opening up" and uses CT to help others?

A one-tailed Kendall's tau correlation analysis was conducted to find out more about the relationship between the months a YP has been working (YPJobMonths) and the reported Mode of Reflection MOReflectionModeYP. A statistically significant positive correlation has been found, $\tau = .529$, $p = .023$, $n = 10$. We may conclude that YPs indeed "open up" more and more in time and use CT to the benefit of others.

In this case it can't be stated for sure what is the cause of the growth of CT Conduct and Impact over time, but it can be argued that a cause may be that, over time, YPs direct their conducting CT more to their outside world. It can be stated that it is not likely that the growth of both CT Conduct and Impact is caused by the growth of "YP feeling the opportunity" to conduct CT.

In order to find out more about the correlation between Impact and Conducting CT, first the group YPs is divided in two subgroups: a group YPs with a higher Impact

score than average, and a group with a lower Impact score than the average Impact score. If Impact is not a phenomenon that has Influence on the Conduct of CT over time, both subgroups should give more or less the same CT Conduct improvement over time.

Pearson correlation tests were conducted. The results showed a non-significant positive correlation between YPJobMonths and MOBehaveTotal with $r(2) = 0.667$, $p = .166$ for the high performing group. Also there was a non-significant positive correlation between YPJobMonths and MOBehaveTotal with $r(1) = 0.858$, $p = .343$ for the low performing group. Given the small number of datapoints we must interpret this numbers with care, but we may not assume that high performance on Impact by YPs correlates with (or "predicts" or "leads to") more conducting CT.

Then the group YPs was divided in two other subgroups: a group YPs with a higher CT Conduct score than average, and a group with a lower CT Conduct score than the average CT Conduct score. If CT Conduct is not a phenomenon that has Influence on the Impact, both subgroups should score more or less the same Impact improvement over time.

Pearson correlation tests were conducted. The results showed a significant and very strong positive correlation between YPJobMonths and MOImpactTotal with $r(3)=0.914$, $p=.030$ for the high CT conducting group, and a non-significant positive correlation between YPJobMonths and MOImpactTotal with $r(3) = 0.519$, $p = .370$ for the low CT conducting group. Given the small number of datapoints we must interpret this numbers with care, but in indicates that high performance on Conducting CT by YPs correlates with (or "predicts" or "leads to") more Impact.

Another idea was to find out more about the relationship between MOImpactOrg and the conduct of CT (MOBehaveTotalYP). If conducting CT was of importance, the YPs that conduct CT will over time, regarding their Impact, "float to the surface". The

group of 10 YPs was divided in a “newbies” group and a “established” group, based on the years they’ve worked (YPJobMonths). Both groups consisted of 5 members. It is to be expected that, when we call a higher than the average Impact score within each group (MOImpactOrg_groupmember>mean(MOImpactOrg_group)) a “success”, the relative number of YPs of “high scoring CT” in the established group that were a success would be higher in the established group, compared to the number found in the newbies group.

It was; there were twice as much high conducting CT YPs in the upper Impact score part in the established group, as there were in the newbies group. However, this sounds more impressive than it is. With group sizes of 5 members, and 1 high CT scoring success in the newbies group and two in the established group, this does not come near to a statistically significant result. Using a one-tailed two-sample proportion Z-test, we can conclude that with 1 success out of 5 possibilities for the newbies group, against 2 successes out of 5 possibilities for the established group, the established group did not score significantly higher than the newbies (p = .490) on “CT conducting members floating to the surface”).

To summarize this statistical analysis of the relationship between Conducting CT and Impact, we may conclude that the level of Conducting CT predicts to a very high degree the level of Impact a YP makes, especially when the reflection mode is “Friendly/supporting” and “Outward oriented”. The level of Making Impact does not predict to the same very high degree the Conduct of CT. Also, there is “circumstantial” non-statistically significant evidence that conducting CT precedes making Impact. Note that “preceding” is not the same as “Causing”. But if there should be chosen a “leverage” to let YPs make more Impact, enhancing outward orientated CT Conducting should be

the first choice. Also, this would be the practical one; it seems easier to support YPs with them conducting CT, than with them making Impact.

The development of the Organizational Impact (the "Mental Model" of the group/organization) and environmental Impact (the Impact outside the organization) over time, is visualized (Figure 13). To do this, every score on both variables MOImpactOrgTotal and MOImpactEnvirTotal were "normalized" up to the % relative to the number of points that could be scored. This way, every score was translated to a number between -and including- 0 and 100. The scores of the YP's were sorted on the variable YPJobMonths. A timeline was made, with a "binsize" of 5 months, and the scores of MOImpactOrgTotal and MOImpactEnvirTotal were plotted on the timeline. The cumulative values were plotted on the same timeline and a graphic was made. Figure 13 shows that, as we saw, there is more Impact made on the organization then on the Environment.

Also it seems that there is a kind of a "timelag" between the Impact on the (Mental Model of the) organization and the Impact on the Environment, suggesting the first precedes the second. But do the "central tendencies" of the time effects take place really differ, as we would expect when effects in the environment "lags" on effects in the organization?

The normalized scores of variables MOImpactOrgTotal and MOImpactEnvirTotal were divided by 10 and "translated in time", making it possible to find the central tendency of the time score per variable (e.g. a normalized score of 20 on MOImpactOrgTotal in YPJobMonths 14, resulted in [20/10=] 2 scores of 14). A row of MonthImpactOrg and a row of MonthImpactEnvir resulted, representing the month Impact took place. A one-tailed Mann-Whitney test for two Independent Samples was

conducted to test if there was a difference between the “MonthImpactOrg” and “MonthImpactEnvir” on Month. The results showed virtually no difference between the groups, $U = 126$, $p = .178$, $n1 = 22$, $n2 = 14$. The median for the MonthImpactOrg group was 58, the MonthImpactEnvir group was 60. This suggests that there is a small but not significant timelag between the reported effects on the Organization and the effect on the Environment, so changes in the Mental Model do not have to proceed changes in the Environment.

On the same data a one-tailed independent T-test was conducted (knowing that the assumption of normality may be invalidated, as a Shapiro-Wilk test rejected both groups as Normal distributed with $p = .003$x, and d'Agostino-Pearson accepted both groups as Normal distributed with $p = .07$ and $p = .08$) to test if there was a significant difference between “MonthImpactOrg” and “MonthImpactEnvir” on Month. The results showed a non-significant difference, $t(34) = .873$, $p = .194$, $d = .298$, The mean score for the MonthImpactOrg group was 44.818, $SD = 26.032$. The mean score for the MonthImpactEnvir group was 52.428, $SD = 24.632$. This suggests that there might be a time lag of about 7 months between the reported effects on the organization and the effect on the environment, so changes in the Mental Model might have to proceed changes in the environment. This is however, again, not statistically significant.

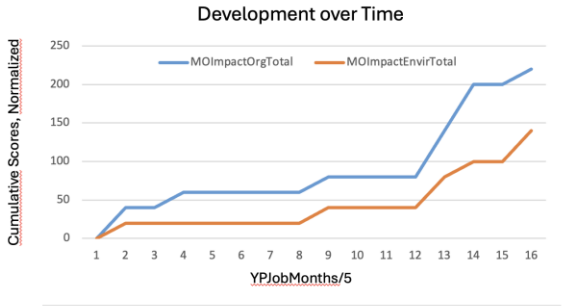


Figure 13:
Development of Impactscores over time (cumulative)

4.3.5 CT OMA and Impact

As found in paragraph 4.3.3 , Conducting CT and Impact correlate. Does the combination of CT Ability, CT Motivation and CT Opportunity correlate with Impact?

First we may conclude that we cannot use Motivation as it is likely that all respondents were motivated (see paragraph 4.5.1), so this variable will be ignored in this paragraph. Opportunity is measured with 3 ordinal variables YPBelonging, YPCTAllowed and YPCTFreedom. CT Ability is measured in a 100-point "ratio" scale with YPChallengeScoreReported. Impact is measured in a calculated ratio scale. There is no "OMA calculation" defined as a whole, Therefore Opportunity (stated as Opportunity, being the sum of YPBelonging, YPCTAllowed and YPCTFreedom) and the CT Ability (YPChallengeScoreReported) are both separate analyzed to find out if they correlate with Impact (MOImpactTotal).

A one-tailed Spearman's rang correlation test was conducted to analyze the relation between Opportunity and MOImpactTotal. The results showed a non-significant weak correlation, $\rho = 0.048$, $p = .445$, $n = 11$. This suggests that there is no relation between the O from OMA and making Impact.

A one-tailed Pearsons's correlation test was conducted to analyze the relationship between YPChallengeScoreReported and MOImpactTotal. Results showed a non-significant correlation, $r = 0{,}367$, $p = .133$, $n = 11$. Combined with the test before, these results suggests that there might be a weak but not significant relationship between CT OMA and Impact. If there should be a correlation, the CT Ability has more weight than Opportunity.

4.4 Relationships between "Miscellaneous constructs", CT and Impact

4.4.1 Teamsize, Conducting CT and Making Impact

Is it easier to Conduct CT in a smaller team? A one-tailed Kendall's tau correlation analysis was conducted to investigate the relationship between YPTeamSize and MOBehaveYPTotal for the 11 paired MOs and YPs. A weak positive statistically non-significant relationship has been found, $\tau = 0.097$, $p = .359$, $n = 11$. This suggests that team size has no influence on conducting CT.

Is it easier to make Impact in smaller teams? A one-tailed Kendall's tau correlation analysis was conducted to investigate the relationship between YPTeamSize and MOImpactTotal for the 11 paired MOs and YPs. A negative but statistically non-significant relationship has been found, $\tau = -0.367$, $p = .087$, $n = 11$. This suggests that it is probable that the bigger the team, the less Impact a YP can make. However, this is not statistically significant.

4.4.2 Branche, Conducting CT and Making Impact

Do YPs show more CT conduct in specific branches? The group was divided in a MOISIC K26 group and a "Other" group. A two-tailed Mann-Whitney test for two Independent Samples was conducted to test if there was a difference between the K26 and the Other group on MOBehaveTotalYP. The results showed a non-significant difference between the groups, $U = 9{,}5$, $p = .412$, $n1 = 7$, $n2 = 4$. The median for the K26 group was 75, the other group was 93. This suggests that there is no difference in conducting CT in a specific Branche.

Do YPs have more Impact in specific branches? The group was divided in a MOISIC K26 group and a "Other" group. A two-tailed Mann-Whitney test for two

Independent Samples was conducted to test if there was a difference between the K26 and the Other group on MOImpactTotal. The results showed a non-significant difference between the groups, $U = 4{,}5$, $p = .072$, $n1 = 7$, $n2 = 4$. The median for the K26 group was 1, the other group was 4.5. This suggests that may be differences in Impact YPs can make in different Branches, but this is not statistically significant.

4.5 "Flat" scores

4.5.1 Motivation

Motivation is not measured in this Survey. Joining the survey (and 19 out of 31 YPs initially wanting to obtain a C10th certificate) is an indication that there is willingness to conduct CT behavior. Also, YPs were given an opportunity to "cheat" with the scores obtained by the challenge. Do they take the easy way and don't take the test and just fill out the scores, do they fill out a higher score than obtained in reality (and try to "look better"), or do they do what someone would do who wanted to "grow"? There was no YPChallengeScoreReported that differed from the YPChallengeScoreGained (filled out by the researcher based on the actual test). It may be concluded that all the participating YPs have at least to some degree a motivation to conduct Critical Thinking. This leads to the observation that there is no point in analyzing the correlating of Motivation with other constructs. That is because all the values of Motivation that would be used would be uniform.

4.5.2 Size of the Organization

Because all but 1 of the YP/MO pairs worked in an organization with the same size (organization category), there is no point in analyzing the correlating of MOOrgCat

with other constructs. That is because nearly all the values of MOOrgCat that would be used would be uniform.

4.6 Summary of the Obtained Results

In this chapter the data of the survey was analyzed. The next conclusions were drawn:

- 44 respondents (31 YP, 12 MO) conducted the survey, leading to 11 matched MO/YP pairs.
- All but 1 MO was manager of the YP.
- 7 of the 11 MO/YP pairs worked in the Branche “Computer programming, consultancy and related activities”, and all but 1 pair worked in a Large SME.
- Half of the YP worked with up to 5 colleagues daily, but no one worked with more than 20 colleagues.
- The education level of the members of both MO and YP groups were more or less equally distributed (half Research University/WO, half applied Sciences University/HBO).
- The YPs were on average 2.6 years ago graduated with a substantial standard deviation of 2.1 years. Seasonal influences were visible in the composition of the group.
- The HEI-way of thinking (CT Ability) was by 81% of the YP and 73% of the MO mentioned as a still-relevant big benefit of their education.

- WO (Research University) educated respondents scored the benefit of their education on the dimension “Knowledge” significantly lower than HBO (Applied Universities) did (p = .028).
- The idea that YP’s made Impact on their environment was around 4 on a five-point scale by both YPs and MOs. However, this score cannot be used as a substitute for the Impact measurement as defined in this Research.
- MOs score the opportunity YPs get to conduct CT, significantly higher than YPs themselves.
- The MOs and YPs give the same score to the extend YPs tend to use CT to challenge others in their thinking (“Reflection mode”). This score was for all but 1, externally directed (3 or more on a five-point scale).
- All the predictions made based on the thinking model were not falsified with the data. Therefore, the thinking Model is not rejected. However, this conclusion is, because of the limited number of respondents, not based on more advanced statistical techniques.
- Impact and CT Ability are not correlated, Conducting CT and Impact are.
- There is a difference in CT Ability between WO and HBO YPs, but not in their CT Conduct.
- Both CT Conduct and Impact YPs make, grow over time. This may be due to the finding that YPs “slowly open up” and direct their CT conduct over time more and more to the benefit of others.
- Causation between CT Conduct and Impact can’t be proved based on a survey-research. It’s possible they influence each other. It is likely that stimulating CT Conduct is a “leverage” to steepen the growth in Impact a YP can make.

- It is possible that it is harder for a YP to make Impact in a bigger team, but a correlation between the two is not proven. Conduction CT by YPs does not correlate with size.
- The Impact an YP makes may probably differ per Branche, the Conduct of CT is not Branche specific.
- All YP that joined the survey seemed motivated to Conduct CT.

This chapter was aimed at “number crunching”. In Chapter V the results from this chapter will be used to answer the research questions and other notable interpretations.

CHAPTER V:
DISCUSSION

In this chapter the results of Chapter IV will be used to answer the sub-questions as stated in paragraph 3.4, and using these questions, answer the first two main research questions. The implications of the findings, and thus the answer on the third main research question what actions HEIs, YPs and Employers can take to enhance the Impact of YPs, will be discussed in Chapter VI

5.1 Sub Questions and Answers

This chapter will answer, based on the results as discussed in Chapter IV, the sub-questions in order as they are stated in Paragraph 3.4.

5.1.1 Operationalization and Quality of the Thinking model

The first question to be answered was “Does the thinking model, regarding the assumptions behind the calculated Impact score, has any explaining or prediction power?”. Well, had it?

Based on the results in Chapter IV, it can be stated that all predictions stemming from the Simplified Thinking Model held (so far). Although the number of responses was too low for a sound statistical underpinning, all expectations stemming from the operationalized Thinking model were not falsified by the real data. The expectations about the relative frequencies of the Impact variables, based on the weights and way of calculating these calculated variables, also did not deviate statistically significant from the measurements in reality.

The developed CT Ability test had a satisfying enough quality (Cronbach's alpha of .644 that could be brought up to .722) within the limits of needed test-taking-time that was stated as acceptable. It was, according to some YPs the researcher spoke afterwards, also "fun" to do the test.

Although it was a bit painful to the researcher, the further simplification of the thinking model also was workable. The biggest sacrifice was to not find out in detail if conducting CT led to changes in the Mental Model of the YP, which could lead to changes on the Mental Model of the team/organizational level, et cetera. This was a pity, because the effect of CT on Mental Models is "undeveloped land". But it would simply take too much time for the respondents to go in such detail and would have probably led to an even lower number of responses. Instead of this level of detail, Impact was seen as "everything outside the YP". The "degree of external orientation of conducting CT" combined with the eight observable CT behaviors in a "from" to "up to" scale resulting in a score for CT Conduct was directly compared with changes in the YP-external world (Impact).

The notice as stated in Chapter II that a "Mental Model" can exist on a personal, group and organizational level made it possible to easily combine these level-specific Mental Model concept into one integrated "Organizational" Impact measurement while keeping the notice of perceiving, feeling and Self. This way the survey could be kept relatively short for this Phenomenon.

Also, the breakup of the changes in the "YP external world", into at one hand the Organizational (Mental Model) Impact and at the other hand the Environmental Impact, seemed to work. As did the translation of the Actor-centric idea about what "Environmental Impact' was into the levels of changes in "horizons". These "horizons" were lent from theories on Innovation frameworks. This Actor-centric idea of defining

Impact was also “new”, compared to the more common specific field of expertise Impact models, as we saw in Chapter II. But also, this seemed an effective choice in this Research project.

The respondents could all in all relate to these choices, given the feedback during the interviews the researcher did on the near-production version with the test-respondents, and the near total absence of the use of “free comment fields” in the final survey.

The researcher therefore states that this thinking model served its purpose (enough) and was workable for this project. But it is surely not a “CT-Impact Theory of Everything”. The way the calculated variables were obtained, and the weights itself are merely “experience and literature based educated guesses” and not sound scientific obtained formulas.

Also, the observation that both Impact and CT Conduct grow over time cannot be fully “explained” by the model (although it may be stated that the degree of external orientation of conducting CT by YP’s also grows in time). The model therefore surely has its limits.

However, using the text from the working definition in 3.2.1, this operationalized model itself served to explain past phenomena, explore present dynamics, and estimate future outcomes. The predictions “held”, the phenomena could be “measured”, and their correlations could be examined. Therefore, it was a useful model with (enough) “power”.

5.1.2 Measured Variables and their Effect on Impact

Using the Thinking Model and the measurements, the next main conclusions can be drawn:

- Having a higher ability to conduct CT does not result into YPs making more Impact. It is the conduct of CT that has a high correlation with Impact.
- Conducting CT and making Impact have an extremely high correlation (for high-Impact high-CT friendly-voiced, external oriented conducting YPs, a statistically significant Pearson's *r* of .914 was observed). If (what could not be tested in this research project) conducting CT precedes making Impact, for this group, the score on CT Conduct predicts 83.5% of the Impact an YP makes.
- Although the evidence is circumstantial and not statistically significant, it is likely that stimulating conducting external oriented CT is an important leverage to making more Impact.
- Over time, YPs don't feel they get more opportunity to conduct CT. However, over time their level of conducting CT and making Impact grows. This may be due to the finding that YPs "slowly open up" and direct their CT conduct over time more and more to the benefit of others. Maybe also their self-confidence grows doing so.
- Managers think the YPs have more opportunity to conduct CT than YPs themselves. It therefore may be wise for managers to pay explicitly and ongoing attention in their contact with YPs to stimulate them to conduct CT, as this conducting CT likely leads to Impact on the phenomena managers deem important.
- Conducting CT can be done in any team size and in any Branche of industry. Making Impact may be more difficult in bigger teams.

- About three out of four respondents (both YPs and MOs) mentioned that the “HEI-way” of thinking was, years after their graduation, still a big benefit of their study. The “specific knowledge of the field of expertise” as a benefit, scored considerably lower for WO (Research University) respondents, but slightly higher for the HBO (Applied Universities) respondents. The benefit “network they obtained during their study” scored lowest. Only the benefit “formal evidence that gave respondents the opportunity to do a job at the desired level” scored higher than the benefit of “thinking in a HEI way”. Impact aside, this fact on itself emphasizes the importance of CT in the HEI curricula.

5.2 Main Research Questions

What are, in short, the answers two the first two main research questions as they were stated in Paragraph 3.4?

- Is there a relationship between Critical Thinking by HEI graduated Young Professionals and to the Impact they make at the start of their careers? Yes, there is a statistically significant relationship between conducting friendly-voiced external oriented CT and making Impact. The higher the MO scores the YP on conducting CT, the Higher the MO scores the YP on making Impact.
- Does Critical Thinking by HEI graduated Young Professionals lead to making Impact at the start of their careers? This causation could not be answered in this research project due the limited time and low numbers of respondents that could attend the training. However, data shows it is likely that conducting friendly-voiced external oriented CT precedes making

Impact. Also, that YP's that conduct this kind of CT also end up in the group that make relatively high Impact. Therefore, stimulating conducting CT would be a logical leverage to enhance the Impact of YPs.

In the coming Chapter VI, the last remaining research question will be answered. "What actions can YPs, HEIs and Employers take regarding learning and conducting CT that will most probably lead to the Impact YPs make at the start of their career?"

CHAPTER VI:
SUMMARY, IMPLICATIONS, AND RECOMMENDATIONS

6.1 Summary

Critical Thinking (CT) is considered as an important competence for the coming decennia to make Impact. However, little research was available if this claim holds.

Based on a literature study a thinking model of the relationship between the CT Opportunity, Motivation and Opportunity, Conducting CT, Mental Models and Impact was proposed (Figure 1). Due to practical boundaries this model was simplified even further (Figure 3). The model was operationalized, including a CT Ability test and a survey. YPs and MOs were found using a diversity of channels, resulting in 31 YPs and 13 MOs answering the survey. 11 MO/YP pairs could be made. A training, aimed at finding out if conducting CT by YPs caused Impact, could not be conducted.

The results of the survey led to the finding that the operationalized thinking model and Ability test were a “good enough” base for this research project. The predictions of the model were not falsified.

It was, based on the survey results, concluded that CT Ability was not strongly correlated to making Impact, but Conducting friendly-voiced external oriented CT was. There is circumstantial evidence that stimulating friendly-voiced external oriented CT Conduct may result in making Impact.

6.2 Implications

The results of this research can be of value for HEIs, Employers of YPs and YPs themselves. In this paragraph some implications are given.

6.2.1 Implications for HEIs

The results from this study may be of value for HEIs, considering the following:

- CT conduct may serve different purposes during the study. In the first year(s), CT Ability should have attention, as it is shown that good CT Ability speeds up the learning curve and enhances problem solving skills (Helsdingen *et al.*, 2010). Also collective Critical Thinking enhances team corporation (Grichanik, 2014). In the later years, when students conduct real-life projects in organizations, the emphasis should be on Conducting CT, maybe as a part of "Advisory skills". This is because it is not the Ability to CT that makes Impact, but the conduct of CT.
- Special attention should be given to let students develop friendly-voiced external oriented CT. This is especially important, as most HEIs only pay attention to short-term CT skills and dispositions (Puig *et al.*, 2019). Various alternatives could be chosen. Although it seems to have more traction in the advisory world than in the scientific community (Passmore and Rowson, 2019), the way "Challenging" the ideas of others in Neuro-Linguistic Programming for example could probably be helpful (Knight, 2012). Also facilitation techniques may come at hand (Turan, Fidan and Yıldıran, 2019).
- Students should be made aware that CT isn't a one-size-fits-all silver bullet. Feelings and opinions of the Self of the student, of his/her colleagues, the group, the organization and the society have a major Impact and are important as well. Not everything can be logically

explained, nor everyone is willing to let his/her Mental Model be smashed by logic. Also not every challenge should a priori be tackled with the same CT-approach (Shao *et al.*, 2022). YPs should have the ability to recognize when it is the time to conduct (external directed) CT, to what extend and when to use additional approaches. It would be wise for HEI's to incorporate in their curricula the lecturing of indicators what "kind of CT" fits what kind of challenge, see Snowden and Boone's Cynefin framework (Savigny, Blanchet and Adam, 2017). It is logical to conduct "Classical" CT (sense, analyze, respond) in the complex-quadrant. CT as tool in action oriented and fundamental research, and the entrepreneurial "effectuation" model (Sarasvathy, 2008), would be logical when confronted with challenges in the complex-quadrant. But, in this Cynefin framework, it is also perfectly logical to choose for Boyd's ODAA loop (Ryder and Downs, 2022) when confronted with a challenge in the "chaos" quadrant, knowing that ODAA and CT are difficult to combine.

- HEI's should decide clearly what the intended outcome for teaching CT is, as this influences the curricula. Is CT a goal in itself, and thus to make "better people" in a philosophical way? Is CT a tool to make "better citizens" and make them more aware of the values of democracy? Is CT a tool to make them able to work on VUCA, "hairy" problems on order to make progress on the Social Development Goals? Is CT a tool to make "better learners" and thus get better didactical results and throughput rates in the institute? Is CT a tool to harden them against "fake news", advertising, big tech, et cetera? Is CT a tool to make them more "suitable"

for the SME's where they should find their first Job? Is it for something else? Answers to these questions will affect the curriculum.

- YPs that have a WO (Research University) background have a higher CT Ability than HBO (Applied University) alumni, but they don't score higher on Conducting CT nor on Impact. It could be an idea to incorporate more practical elements in the WO curricula, as well as elements of how to conduct CT with an externally oriented Friendly Voice. This however would raise other questions, as it would question one of the fundaments of what Research Universities should be (Praamsma, 2006).

6.2.2 Implications for SMEs

The results from this study may be of value for SME, considering the following:

- YPs may be, at the start of their career, be offered "developing trajectories". This trajectories often aim at developing the YP as fast as possible to understand corporate culture, developing technical depth and social skills. (Lau, 2019)(Varrella *et al.*, 2016). Mentoring is often a part of that (Franko and Myttseva, 2022)(Ovalle *et al.*, 2021). These programs try, in a way, to make YPs as fast as possible "organization members", knowing what the experts know and thinking the way the other organizational members do. CT can be a tool to steepen the learning curve. But is fostering the challenging (in a supporting voice) of the status quo, questioning old beliefs and creating awareness of one's assumptions not what YPs could bring also? Are these classical development programs maybe too much of a one-way sending exercise, and is the possible contribution that YPs can bring properly valued? Mentoring can be of

value for the mentor too, when the YP can conduct CT on the ideas and assumptions of the mentor. Challenging old ideas in a center of power may enhance the Impact YPs make.

- In addition to the former bullet, sending might not be the most effective way to change Mental Models, also not of the Mental Model of the YP. To "walk the words" as author of this document….. Given that information easily gets neglected, distorted, enriched or misinterpreted in order to fit into one's Mental Model, what would make more Impact: asking someone the right questions or telling someone the "truth"? And is sending either way preferable above trying to come to a two-way alignment of Mental Models? Also, MO asking questions to the YP may help close the discovered gap between the space MOs think the YP's feel to conduct CT, and the space YPs feel in reality.
- The measurement of Impact, offered in this research, can be used to measure the effect of development trajectories for YPs. Development trajectories cost time and money, so it would be nice to get an idea of the effect of the program and maybe compare different alternatives. With which program does the YP's make the most Impact for the SME?

6.2.3 Implications for YPs

The results from this study may be of value for YPs, considering the following:

- We may assume that conducting CT is more than being able to conduct CT. The ability to conduct CT is, as this study shows, merely a prerequisite for making Impact. Enhancing the Ability for CT is a good start. But it would be wise for YPs, in the first years of your career, to

enhance competencies necessary for understanding dispositions of other Actors and gaining/earning their trust (Friendly Voice) and to a two-way alignment of Mental Models (external oriented CT). This way a YP can improve his/her Impact dramatically. Try not to aim purely on development paths that result in more technical skills and knowledge but incorporate "soft skill" training also. And, try to keep a growth mindset during your daily work.

6.3 Recommendations for Future Research

There are different questions raised stemming from this research:

- What exactly makes that CT Conduct and Impact of YPs grow over time? Is it really the "opening up" of YPs? The Thinking Model does not explain this phenomenon fully. More detailed insights in the factors that lead to growth may show other "leverages".
- The research shows that is likely that conducting friendly-voiced external oriented CT leads to changes in the Mental Models of others. However, due to the limitations in this research, the exact mechanism how Conducting CT leads to changes in one's own Mental Model is not clear in detail. Does conducting CT indeed lead to changes in the perceiving/conscious quadrant, and to what extend does that affect the other quadrants and the "Self"? And what is the exact nature of the connection and influence between the Mental Model "layers" (individual, group/organization)?
- The "weights" in the thinking model are "educated guesses" and seem to work, as shown in this thesis. But is it possible to gain more data and

perform a thorough confirmative factor analysis in order to find out if this "weights", or maybe the complete formula, can be optimized?

- It is likely that CT and Impact have a (yet not completely understood) mutual, circular relationship. It is also likely that stimulating the conduct of friendly-voiced external oriented CT is a good leverage to speed up this circular process. However, it would be nice to find out if this is really the case. The researcher of this thesis has, no longer bound by time-constraints, planned a follow-up to test this assumption.
- When we stimulate CT as stated, does it only speed up the process (YPs make Impact earlier in their career) and is there a "ceiling" in the amount of Impact, or does it also lead to a higher "Impact ceiling". If the latter is the case, stimulating the CT conduct would also be a good idea for professionals later in their career. Further research to find out if the "ceiling" idea, or "the sky is the limit" idea (or maybe a complete other mechanism) holds, could therefore be valuable.
- The "Mental Model"-idea used in this research, based on the work of Ledoux (LeDoux, 2020), has a dimension that is not addressed firmly in this research; Feelings. As stated, one of the most important things is "keeping the Mental Model consistent" (see paragraph 2.1.2). Feelings and the idea of the Self have most likely a huge effect on the "filters", making that sensed signals are omitted, altered and enriched before they are even processed. In the educational trajectories, in order to "grow" YPs out of students, the "perceiving" dimension is often well-addressed. It might be fruitful to start research on how students can be made aware of their feelings and image of the Self. And especially on how this influences the

things they “know” and to what extend this leads to more flexibility in updating their Mental Models. More flexibility in updating Mental Models would possibly lead to Mental Models that reflect reality better and, probably, lead to more Impact (and more resistance against propaganda). This might also be a research direction.

- In addition to the in this research omitted “Feeling” dimension, the quadrant “unconscious perceiving” hasn’t had a lot of attention. In the work of Vlerick an interesting idea was proposed: use “heuristic reasoning” deliberately when confronted with “ancient problems” people dealt with for millions of years (Vlerick, 2022). This system was made for that. Is it wise to ignore un unpleasant feeling stemming from someone that is approaching you in a dark street, just because you can’t explain logically what and why? In the Dutch healthcare this mix of “Feeling” and “Heuristic thinking” is done when triaging people that call the doctors (*Niet-pluisgevoel: een diagnostisch instrument | H&W*, no date). It is concluded that this instrument is unprecise, but very useful. In this research a diversity of callers that would have been only “helped” by telephone using the models stemming from the “conscious perceiving” quadrant, but where the triagist had such a feeling, were discussed with the doctor. It almost always led to follow-up research. In two-third of the cases it led to the diagnose “cancer”. In short: it might be useful to do more research on when to deliberately use which type of thinking and incorporating these insights in the students’ curriculum.

6.4 Remarks on the Validity and Reliability of the Research

There are some remarks to be made on this research:

- The absence of the causation test. There is in this research project no test conducted that aimed on stimulating CT by YPs, and a measurement afterwards what the relative changes in reported Impact was, in relation to a “baseline” Impact. There can’t be said anything definitive about conducting CT causing effects on Impact. Although this is true, in paragraph 4.3.4 a variety of test were conducted to be able to find out a bit more of the nature of the relationship between conducting CT and Impact.
- Calculated variables are treated as "ratio" variables in the statistical tests. The calculated variables were built up out of ordinal variables and “weights”. The idea about this using “weights” was to make the “distance” between the scores meaningful, and that a score of 0 indeed means the absence of the measured construct (EG: a 0 score on MOImpactEnvirYP means “No Impact on the environment” and a score of 2 compared to a 1 as “twice as much Impact”. This even while the units/dimensions of measurement of “Impact” are vague). However, there might be different ideas about treating calculated variables as Ratio variables, as all the scores of all calculated variables don’t have a dimension. Therefore, in this project sometimes also variations of proportion tests were used.
- In some cases, assumptions and the usage of statistical tests is not standard (for example, the chi-square test to check the weights of the Simplified Thinking Model). Some may find this creative and unconventional; others may judge it as questionable. Retesting and higher response numbers

would enhance the reliability considerably, and also higher response rates would open the way for other ways of testing.

- The use of a survey. Surveys are not the most reliable data source (Tourangeau, 2021)(GALESIC and BOSNJAK, 2009). It can't be ruled out that the interpretation of questions as well as the sampling method introduced structural flaws in the dataset. The researcher tried to test the tools with validity and reliability in mind, using not only statistics but also interviews in the test rounds (triangulation).
- The use of self-build measurement tools. When possible, it is to the opinion of the researcher wise to use existing tools, as this prevents the researcher from pitfalls and problems others encountered and overcame. However, the researcher did not find existing tools that measured the phenomena that should be measured. Especially no tools that were within budget and didn't consume a lot of time for the respondents. Using self-developed tools that cannot be validated against other tools may introduce the risk of standard biases in measurements. Again, the researcher tried to test the tools with validity and reliability in mind, using not only statistics but also interviews in the test rounds (triangulation).
- The way respondents were found. The pre-information about the research project was something like "The Impact YPs make with the HEI-way of thinking". Using this introduction, it might not be of a great surprise that a significant number of respondents in the survey indicate that the HEI-way of thinking was important to them. That is a bit like first placing a big flashy sign "Fresh Fried Tasty Fish" in front of a delicious-smelling seafood stall and then asking the people that arrive if they appreciate fried

seafood. However, in this case the question was also put to others (the MOs), and in this case their answers were in line with the answer of the YPs that were attracted to this research-seafood stall. This should somewhat dampen this “fried fish phenomenon”.

- The first idea of the researcher was to aim at YPs that would not have worked for more than 3 years, as this would eliminate the possible “feedbackloop” of “Impact” on “CT Conduct”. Also, the research was aimed at alumni from the Rotterdam University of Applied Sciences (RUAS), especially the Rotterdam Business School (RBS). Due to the problems as earlier described, the definition of YP was stretched to a time period of 8 years, and the pool became as big as all the alumni from the HEIs in the Netherlands. Of course, this diminished the usability of the results a bit for the RUAS. However, it increased the usability for the other HEIs in the Netherlands. Extrapolation of results (e.g. for all students outside the Netherlands) is always risky.

It is wise to keep these limitations in mind when using the results of this research.

6.5 Conclusion

There is strong evidence that Conducting Friendly-voiced external oriented CT by Young Professionals and Making Impact are strongly correlated. Stimulating this kind of conduct of CT by YPs is probably an effective leverage to enhance, or at least speed up the process of, making Impact by YPs at the start of their career.

REFERENCES

Andresen, B.B. (2019) 'Sustainable E-Learning–A Case Study on the Pros and Cons of Certification', in. *Sustainable ICT, Education and Learning: IFIP WG 3.4 International Conference, SUZA 2019, Zanzibar, Tanzania, April 25–27, 2019, Revised Selected Papers 1*, Springer, pp. 30–38.

Arain, M. *et al.* (2013) 'Maturation of the adolescent brain', *Neuropsychiatric Disease and Treatment*, 9, pp. 449–461. Available at: https://doi.org/10.2147/NDT.S39776.

Arieli, S., Sagiv, L. and Roccas, S. (2020) 'Values at Work: The Impact of Personal Values in Organisations', *Applied Psychology: An International Review*, 69(2), pp. 230–275. Available at: https://doi.org/10.1111/apps.12181.

Balasubramanian, R.B.R. and Fathima, M.P. (2011) 'Thinking and Decision Making - An Overview', *Indian journal of applied research*, 3, pp. 185–186.

Beck, A.T. (2005) 'The current state of cognitive therapy: a 40-year retrospective: Archives of general psychiatry', *Archives of general psychiatry*, 62(9), pp. 953–959. Available at: https://doi.org/10.1001/archpsyc.62.9.953.

Beck, D.E. and Cowan, C.C. (2014) *Spiral Dynamics: Mastering Values, Leadership and Change*. John Wiley & Sons.

Beck, K. (2019) 'On the Relationship Between "Education" and "Critical Thinking"',

Frontiers and Advances in Positive Learning in the Age of InformaTiOn (PLATO) [Preprint]. Available at: https://api.semanticscholar.org/CorpusID:212853302.

Becker, S.D. and Endenich, C. (2023) 'Entrepreneurial Ecosystems as Amplifiers of the Lean Startup Philosophy: Management Control Practices in Earliest-Stage Startups', *Contemporary Accounting Research*, 40(1), pp. 624–667. Available at: https://doi.org/10.1111/1911-3846.12806.

Bertsch, A. (2011) 'Representing the Concept of Culture: Has the Time Come to Replace the Layered Onion?', *Oxford Journal*, pp. 19–26.

Burris, E.R. (2012) 'The Risks and Rewards of Speaking up: Managerial Responses to Employee Voice', *The Academy of Management Journal*, 55(4), pp. 851–875.

Cambridge Dictionary, none (2023) *impact*. Available at: https://dictionary.cambridge.org/dictionary/english/impact (Accessed: 31 October 2023).

Cannon, H.M., Feinstein, A.H. and Friesen, D.P. (2010) 'Managing Complexity: Applying the Conscious-Competence Model to Experiential Learning', *Developments in Business Simulation and Experiential Learning: Proceedings of the Annual ABSEL conference*, 37. Available at: https://absel-ojs-ttu.tdl.org/absel/article/view/306 (Accessed: 6 March 2024).

Catchings, G. (2015) 'A Practical Coaching Model for Critical Thinking Skill and Leadership Development (C/CTSLD)', *Management and Organizational Studies*, 2(4), p. p42. Available at: https://doi.org/10.5430/mos.v2n4p42.

Chao, Y.-L. and Lam, S.-P. (2011) 'Measuring Responsible Environmental Behavior:

Self-Reported and Other-Reported Measures and Their Differences in Testing a Behavioral Model', *Environment and Behavior*, 43(1), pp. 53–71. Available at: https://doi.org/10.1177/0013916509350849.

Cline, B.J. (2016) 'A Psychological Account of the Formation of Self Deceptive Beliefs', in. Available at: https://api.semanticscholar.org/CorpusID:253201593.

Cohen, M. (2015) *Critical Thinking Skills For Dummies*. John Wiley & Sons.

Cohen, M.S. (2000) 'A three-part theory of critical thinking: Dialogue, mental models, and reliability', in. *4. TITLE AND SUBTITLE Training Critical Thinking Skills for Battle Command: ARI Workshop Proceedings*, p. 50.

Covey, S.R. (2020) *The 7 Habits of Highly Effective People: 30th Anniversary Edition*. Simon and Schuster.

D'Alessio, F.A., Avolio, B.E. and Charles, V. (2019) 'Studying the impact of critical thinking on the academic performance of executive MBA students', *Thinking Skills and Creativity*, 31, pp. 275–283. Available at: https://doi.org/10.1016/j.tsc.2019.02.002.

De Bie, H., Wilhelm, P. and Van Der Meij, H. (2015) 'The Halpern Critical Thinking Assessment: Toward a Dutch appraisal of critical thinking', *Thinking Skills and Creativity*, 17, pp. 33–44. Available at: https://doi.org/10.1016/j.tsc.2015.04.001.

Devassy, S.M. *et al.* (2023) 'Development of immersive learning framework (ILF) in achieving the goals of higher education: measuring the impact using a pre–post design', *Scientific Reports*, 13(1), pp. 1–9. Available at: https://doi.org/10.1038/s41598-023-45035-0.

Dharani, B. and April, K. (2022) ‘Understanding and Reformulating Mental Models’, *Innovative Leadership in Times of Compelling Changes: Strategies, Reflections and Tools*, pp. 155–169.

Domínguez, E. *et al.* (2019) ‘A taxonomy for key performance indicators management’, *Computer Standards & Interfaces*, 64, pp. 24–40. Available at: https://doi.org/10.1016/j.csi.2018.12.001.

Dungan, J., Waytz, A. and Young, L. (2015) ‘The psychology of whistleblowing’, *Current Opinion in Psychology*, 6, pp. 129–133. Available at: https://doi.org/10.1016/j.copsyc.2015.07.005.

Dunning, D., Heath, C. and Suls, J.M. (2004) ‘Flawed Self-Assessment: Implications for Health, Education, and the Workplace’, *Psychological Science in the Public Interest*, 5(3), pp. 69–106. Available at: https://doi.org/10.1111/j.1529-1006.2004.00018.x.

edu.nl - Dé URL-shortener voor onderwijs en onderzoek met respect voor privacy. (no date). Available at: https://edu.nl/ (Accessed: 18 March 2025).

EFDM (2023) ‘BSIS Assessment Criteria Guide: The comprehensive multi-dimensional impact assesment tool for business school’. EFMD Global Network.

Errida, A. and Lotfi, B. (2021) *The determinants of organizational change management success: Literature review and case study*. Available at: https://doi.org/10.1177/18479790211016273.

Franko, N. and Myttseva, O. (2022) ‘MENTORING IN IT THE WAY TO IMPROVE THE SKILLS OF A YOUNG SPECIALIST’, in *Theoretical and Applied Aspects*

of Device Development on Microcontrollers and FPGAs 2022. *Theoretical and Applied Aspects of Device Development on Microcontrollers and FPGAs 2022*, NURE, pp. 27–28. Available at: https://doi.org/10.35598/mcfpga.2022.009.

Freudenburg, W.R. (2019) 'Social Impact Assessment', *Rural Society in the U.S.: Issues for the 1980s* [Preprint]. Available at: https://api.semanticscholar.org/CorpusID:145369100.

GALESIC, M. and BOSNJAK, M. (2009) 'EFFECTS OF QUESTIONNAIRE LENGTH ON PARTICIPATION AND INDICATORS OF RESPONSE QUALITY IN A WEB SURVEY.', *Public Opinion Quarterly*, 73(2), pp. 349–360. Available at: https://doi.org/10.1093/poq/nfp031.

Gemeente Rotterdam in cijfers en grafieken (bijgewerkt 2024!) (2024) *AlleCijfers.nl*. Available at: https://allecijfers.nl/gemeente/rotterdam/ (Accessed: 22 June 2024).

Global scale - Table 1 (CEFR 3.3): Common Reference levels - Common European Framework of Reference for Languages (CEFR) - www.coe.int (no date) *Common European Framework of Reference for Languages (CEFR)*. Available at: https://www.coe.int/en/web/common-european-framework-reference-languages/table-1-cefr-3.3-common-reference-levels-global-scale (Accessed: 25 February 2025).

Gould, S.J. and Lloyd, E.A. (1999) 'Individuality and adaptation across levels of selection: How shall we name and generalize the unit of Darwinism?', *Proceedings of the National Academy of Sciences*, 96(21), pp. 11904–11909. Available at: https://doi.org/10.1073/pnas.96.21.11904.

Grichanik, M. (2014) ‘The Effects of Collaborative Critical Thinking Training on Trust Development and Effectiveness in Virtual Teams’.

Gudmundsson, H., Joumard, R. and Aschemann, R. (2010) ‘Indicators and their functions’, in *Indicators of environmental sustainability in transport*. Bron: Institut national de recherche sur les transports et leur sécurité – INRETS, pp. 23–43.

Haase, F.-A. (2010) ‘CATEGORIES OF CRITICAL THINKING IN INFORMATION MANAGEMENT. A STUDY OF CRİTİCAL THİNKİNG İN DECİSİON MAKİNG PROCESSES’, *Nómadas. Revista Crítica de Ciencias Sociales y Jurídicas* [Preprint].

Hatcher, D. (2013) ‘The Halpern Critical Thinking Assessment: A Review’, *Inquiry: Critical Thinking Across the Disciplines*, 28, pp. 18–23. Available at: https://doi.org/10.5840/inquiryct201328315.

Heijltjes, A. (no date) *Docentenhandreiking Kritisch Leren Denken*. Available at: https://kritischdenkenhbo.nl/content/read?id=441 (Accessed: 24 June 2024).

Heink, U. and Kowarik, I. (2010) ‘What are indicators? On the definition of indicators in ecology and environmental planning’, *Ecological Indicators*, 10(3), pp. 584–593. Available at: https://doi.org/10.1016/j.ecolind.2009.09.009.

Helsdingen, A.S. *et al.* (2010) ‘The Effects of Critical Thinking Instruction on Training Complex Decision Making’, *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 52(4), pp. 537–545. Available at: https://doi.org/10.1177/0018720810377069.

Hoffman, B.G. (2017) *Red Teaming: Transform Your Business by Thinking Like the Enemy*. London: Piatkus.

Hogeschool Rotterdam (2015) *Hogeschool Rotterdam Business School: introductie*, *Hogeschool Rotterdam*. Available at: https://www.hogeschoolrotterdam.nl/samenwerking/instituten/hogeschool-rotterdam-business-school/Rotterdam_business_school/ (Accessed: 5 September 2023).

Home (no date) *Fibonicci*. Available at: https://www.fibonicci.com/ (Accessed: 26 February 2025).

Home Page (Welcome) | Real Statistics Using Excel (no date). Available at: https://real-statistics.com/ (Accessed: 1 March 2025).

'How Does Theory of Change Work?' (no date) *Theory of Change Community*. Available at: https://www.theoryofchange.org/what-is-theory-of-change/how-does-theory-of-change-work/ (Accessed: 24 October 2023).

Irwin, W. and Bassham, G. (2003) 'Depression, Informal Fallacies, and Cognitive Therapy: The Critical Thinking Cure?', *Inquiry: Critical Thinking Across the Disciplines*, 21(3), pp. 15–21. Available at: https://doi.org/10.5840/inquiryctnews200321335.

jamovi - open statistical software for the desktop and cloud (no date). Available at: https://www.jamovi.org/ (Accessed: 1 March 2025).

Jin, Z. *et al.* (2022) 'Logical Fallacy Detection'. arXiv. Available at: http://arxiv.org/abs/2202.13758 (Accessed: 3 December 2023).

Jones, N. *et al.* (2011) 'Mental Models: An Interdisciplinary Synthesis of Theory and Methods', *Ecology and Society*, 16(1). Available at: https://doi.org/10.5751/ES-03802-160146.

Jordaan, B. (2019) 'Leading organisations in turbulent times: Towards a different mental model', *Leading in a VUCA World*, 38(4), p. 59.

Kahan, D.M. *et al.* (2013) 'Motivated Numeracy and Enlightened Self-Government'. Rochester, NY. Available at: https://doi.org/10.2139/ssrn.2319992.

Kahneman, D. 1934- (2013) *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kansen op de arbeidsmarkt voor afgestudeerde hbo'ers (no date) *HBO Monitor*. Available at: https://www.hbomonitor.nl/nl/resultaten/kansen-op-de-arbeidsmarkt (Accessed: 22 June 2024).

Kellner, A., Cafferkey, K. and Townsend, K. (2019) 'Ability, Motivation and Opportunity theory: a formula for employee performance?', in K. Townsend et al. (eds) *Elgar Introduction to Theories of Human Resources and Employment Relations*. Edward Elgar Publishing. Available at: https://doi.org/10.4337/9781786439017.00029.

Kim, Y.-H. *et al.* (2016) 'Why Do People Overestimate or Underestimate Their Abilities? A Cross-Culturally Valid Model of Cognitive and Motivational Processes in Self-Assessment Biases', *Journal of Cross-Cultural Psychology*, 47(9), pp. 1201–1216. Available at: https://doi.org/10.1177/0022022116661243.

Knapp, P. and Beck, A.T. (2008) 'Cognitive therapy: foundations, conceptual models,

applications and research', *Brazilian Journal of Psychiatry*, 30, pp. s54–s64. Available at: https://doi.org/10.1590/S1516-44462008000600002.

Knight, J. (2012) 'Deletion, distortion and data collection: the application of the Neuro-linguistic Progamming (NLP) meta-model in qualitative interviews', *Australasian Journal of Market & Social Research*, 20(1), pp. 15–21.

Kolesnikova, V.V. (2019) 'FEATURES OF ADAPTATION OF YOUNG PROFESSIONALS TO PROFESSIONAL ACTIVITY', *Expert: Paradigm of Law and Public Administration*, 5(3), pp. 235–246. Available at: https://doi.org/10.32689/2617-9660-2019-3(5)-235-246.

Krejci, J.W. (1997) 'Imagery: Stimulating critical thinking by exploring mental models', *Journal of Nursing Education*, 36(10), pp. 482–484.

Kruger, J. and Dunning, D. (1999) 'Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments', *Journal of Personality and Social Psychology*, 77(6), pp. 1121–1134. Available at: https://doi.org/10.1037/0022-3514.77.6.1121.

Lai, E.R. (2011) 'Critical thinking: A literature review', *Pearson's Research Reports*, 6(1), pp. 40–41.

Lau, H.C. (2019) 'Career Development Essentials for Young E&P Technical Professionals', in *SPE Annual Technical Conference and Exhibition*. *SPE Annual Technical Conference and Exhibition*, Calgary, Alberta, Canada: SPE, p. D021S027R003. Available at: https://doi.org/10.2118/196027-MS.

LeDoux, J.E. (2020) 'Thoughtful feelings', *Current Biology*, 30(11), pp. R619–R623.

Available at: https://doi.org/10.1016/j.cub.2020.04.012.

Lewis, A. and Smith, D. (1993) 'Defining higher order thinking: Theory Into Practice', *Theory Into Practice*, 32(3), p. 131. Available at: https://doi.org/10.1080/00405849309543588.

LinkedIn Analytics (no date). Available at: https://www.linkedin.com/analytics/post-summary/urn:li:activity:7269664919820591104/ (Accessed: 1 March 2025).

'List of cognitive biases' (2025) *Wikipedia*. Available at: https://en.wikipedia.org/w/index.php?title=List_of_cognitive_biases&oldid=1280478334 (Accessed: 17 March 2025).

Mercier, H. and Sperber, D. (2011) 'Why do humans reason? Arguments for an argumentative theory', *Behavioral and Brain Sciences*, 34(2), pp. 57–74. Available at: https://doi.org/10.1017/S0140525X10000968.

Midgley, G. and Rajagopalan, R. (2020) 'Critical Systems Thinking, Systemic Intervention, and Beyond', in G.S. Metcalf, K. Kijima, and H. Deguchi (eds) *Handbook of Systems Sciences*. Singapore: Springer, pp. 1–51. Available at: https://doi.org/10.1007/978-981-13-0370-8_7-1.

Mkb-toets (no date) *RVO.nl*. Available at: https://www.rvo.nl/onderwerpen/subsidiespelregels/ez/mkb-toets (Accessed: 25 February 2025).

Natale, S. and Ricci, F. (2006) 'Critical thinking in organizations', *Team Performance Management*, 12(7/8), pp. 272–277. Available at: https://doi.org/10.1108/13527590610711822.

n.b. (2020) ‘The Impact Path – Het Impactpad’. Avans, Social Enterprise NL and Impact Centre Erasmus. Available at: https://impactpad.nl/english/ (Accessed: 19 March 2024).

Nentl, N. and Zietlow, R. (2008) ‘Using Bloom’s Taxonomy to Teach Critical Thinking Skills to Business Students’, *College & Undergraduate Libraries*, 15(1–2), pp. 159–172. Available at: https://doi.org/10.1080/10691310802177135.

Niet-pluisgevoel: een diagnostisch instrument | H&W (no date). Available at: https://www.henw.org/artikelen/niet-pluisgevoel-een-diagnostisch-instrument (Accessed: 7 March 2025).

Nikolopoulou, K. (2023) *Logical Fallacies | Definition, Types, List & Examples*, *Scribbr*. Available at: https://www.scribbr.com/fallacies/logical-fallacy/ (Accessed: 3 December 2023).

Nygaard, S., Kokholm, A.R. and Huulgaard, R.D. (2022) ‘Incorporating the sustainable development goals in small- to medium-sized enterprises’, *Journal of Urban Ecology*, 8(1), p. juac022. Available at: https://doi.org/10.1093/jue/juac022.

OECD (2021) *Applying Evaluation Criteria Thoughtfully*. OECD. Available at: https://doi.org/10.1787/543e84ed-en.

Ovalle, O. *et al.* (2021) ‘Development of a Mentoring Program as a Strategy for the Empowerment of Young Professionals in the Strengthening of Their Soft Skills’, in *SPE Annual Technical Conference and Exhibition*. *SPE Annual Technical Conference and Exhibition*, Dubai, UAE: SPE, p. D031S041R001. Available at: https://doi.org/10.2118/206217-MS.

Passmore, J. and Rowson, T. (2019) ‘Neuro-linguistic programming: A review of NLP research and the application of NLP in coaching’, 14(1), p. 14.

Phan, HuyP. (2011) ‘Deep Processing Strategies and Critical Thinking: Developmental Trajectories Using Latent Growth Analyses’, *Journal of Educational Research*, 104(4), pp. 283–294. Available at: https://doi.org/10.1080/00220671003739382.

Polivy, J. (2001) ‘The false hope syndrome: unrealistic expectations of self-change.’, *International Journal of Obesity & Related Metabolic Disorders*, 25, p. S80. Available at: https://doi.org/10.1038/sj.ijo.0801705.

Praamsma, J.M. (2006) ‘De universiteit als vormingsinstituut / Deel I: Een traditie van eeuwen; Deel II: Naar een moderne universiteit.’

Puig, B. *et al.* (2019) ‘A systematic review on critical thinking intervention studies in higher education across professional fields’, *Studies in Higher Education*, 44(5), pp. 860–869. Available at: https://doi.org/10.1080/03075079.2019.1586333.

RC MacCallum *et al.* (1999) ‘Sample size in factor analysis: The role of model error’, *Psychological Methods, 4*, 1, pp. 84–99. Available at: https://doi.org/10.1037/1082-989X.4.1.84.

Red Teaming Handbook (2021) *GOV.UK*. Available at: https://www.gov.uk/government/publications/a-guide-to-red-teaming (Accessed: 9 March 2025).

Richmond, B. (2010) ‘The thinking in systems thinking: eight critical skills’, *Tracing Connections: Voices of Systems Thinkers*, pp. 3–21.

Rodrigues, R.L. (2019) ‘edX participants’ profile: analysis of the factors that lead to the

search for certification', *Revista Novas Tecnologias na Educação*, 17(1), pp. 72–81.

Rotilă, V. (2022) 'The Absence of Critical Thinking Skills and its Effects. Case Study: Vaccine Hesitation: Romanian Journal for Multidimensional Education / Revista Românească pentru Educaţie Multidimensională', *Romanian Journal for Multidimensional Education / Revista Românească pentru Educaţie Multidimensională*, 14(3), pp. 1–17. Available at: https://doi.org/10.18662/rrem/14.3/594.

Ryder, M. and Downs, C. (2022) 'Rethinking reflective practice: John Boyd's OODA loop as an alternative to Kolb', *The International Journal of Management Education*, 20(3), p. 100703. Available at: https://doi.org/10.1016/j.ijme.2022.100703.

Sandra P.A. Robinson and Verna Knight (2019) *Handbook of Research on Critical Thinking and Teacher Education Pedagogy*. Hershey, PA: Information Science Reference (Advances in Higher Education and Professional Development (AHEPD) Book Series). Available at: https://hr.idm.oclc.org/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,cookie,uid&db=nlebk&AN=2106018&site=ehost-live&scope=site (Accessed: 26 August 2023).

Sarasvathy, S.D. (2008) *Effectuation: Elements of Entrepreneurial Expertise*. Cheltenham, UK: Edward Elgar Publishing. Available at: https://doi.org/10.4337/9781848440197.

Savigny, D. de, Blanchet, K. and Adam, T. (2017) *Applied Systems Thinking for Health Systems Research: a Methodological Handbook*. McGraw-Hill Education (UK).

'Selection of Indicators' (no date) *DataForImpactProject*. Available at: https://www.data4impactproject.org/prh/overview/selection-of-indicators/ (Accessed: 24 October 2023).

Senge, P.M. (2014) *The Fifth Discipline Fieldbook: Strategies and Tools for Building a Learning Organization*. Crown.

Shao, P. *et al.* (2022) 'Problem-Solving in Product Innovation Based on the Cynefin Framework-Aided TRIZ', *Applied Sciences*, 12(9). Available at: https://doi.org/10.3390/app12094157.

Sinha, D. and Sinha, S. (2020) 'Managing in a VUCA World: Possibilities and Pitfalls', *Journal of Technology Management for Growing Economies*, 11(1). Available at: https://doi.org/10.15415/jtmge.2020.111003.

Skrobach, N. *et al.* (2020) 'Youth Professional Development', *Archive of Clinical Medicine*, 26(2). Available at: https://doi.org/10.21802/acm.2020.2.8.

Statistiek, C.B. voor de (2023) *Bijna 6 op 10 baanwisselaars korter dan twee jaar in dienst*, *Centraal Bureau voor de Statistiek*. Available at: https://www.cbs.nl/nl-nl/nieuws/2023/08/bijna-6-op-10-baanwisselaars-korter-dan-twee-jaar-in-dienst (Accessed: 26 June 2024).

Stojanović, B. (2013) 'The Riddle of Thinking Thinking, Fast and Slow by Daniel Kahneman', *Panoeconomicus*, 60(4), pp. 569–576.

Synoniemen van kritiek; ander woord voor kritiek (no date) *Synoniemen.net*. Available at:

https://synoniemen.net/index.php?zoekterm=kritiek (Accessed: 27 June 2024).

Tadema, J. (2020) 'Niveaubeschrijvingen NLQF per descriptor met toelichting'.

Tavakol, M. and Dennick, R. (2011) 'Making sense of Cronbach's alpha', *International Journal of Medical Education*, 2, pp. 53–55. Available at: https://doi.org/10.5116/ijme.4dfb.8dfd.

Thompson, L. and Van Boven, L. (2003) 'A Look into the Mind of the Negotiator: Mental Models in Negotiation'. Rochester, NY. Available at: https://papers.ssrn.com/abstract=1532577 (Accessed: 21 November 2023).

Thonney, T. and Montgomery, J.C. (2019) 'Defining Critical Thinking Across Disciplines: An Analysis of Community College Faculty Perspectives: College Teaching', *College Teaching*, 67(3), pp. 169–176. Available at: https://doi.org/10.1080/87567555.2019.1579700.

Tourangeau, R. (2021) 'Survey Reliability: Models, Methods, and Findings', *Journal of Survey Statistics and Methodology*, 9(5), pp. 961–991. Available at: https://doi.org/10.1093/jssam/smaa021.

Turan, U., Fidan, Y. and Yıldıran, C. (2019) 'Critical Thinking as a Qualified Decision-Making Tool', *Journal of History, Culture & Art Research / Tarih Kültür ve Sanat Arastirmalari Dergisi*, 8(4), pp. 1–18. Available at: https://doi.org/10.7596/taksad.v8i4.2316.

UNSD — ISIC (no date). Available at: https://unstats.un.org/unsd/classifications/Econ/isic (Accessed: 22 February 2025).

Uribe-Enciso, O.L., Uribe-Enciso, D.S. and Vargas-Daza, M.D.P. (2017) 'Critical thinking and its importance in education: some reflections', *Rastros Rostros*, 19, pp. 78–88.

Varrella, G. *et al.* (2016) 'Growing Our Own: A Longitudinal Evaluation of a Professional Development Program for Early-Career 4-H Professionals', *Journal of Extension*, 54(5). Available at: https://doi.org/10.34068/joe.54.05.22.

Vazire, S. and Mehl, M.R. (2008) 'Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior', *Journal of Personality and Social Psychology*, 95(5), pp. 1202–1216. Available at: https://doi.org/10.1037/a0013314.

van de Ven, M. *et al.* (2023) 'Key performance indicators for business models: a systematic review and catalog', *Information Systems and e-Business Management*, 21(3), pp. 753–794. Available at: https://doi.org/10.1007/s10257-023-00650-2.

Versterken van de regio kan niet zonder duurzaam investeren in regionale spilfunctie hogescholen (no date). Available at: https://www.hbo.nl/actueel/actualiteiten/versterken-van-de-regio-kan-niet-zonder-duurzaam-investeren-in-regionale-spilfunctie-hogescholen (Accessed: 22 June 2024).

Vlerick, M. (2022) 'Kritisch en wetenschappelijk denken: een korte introductie'. Available at: https://doi.org/10.26116/WA61-M074.

Why does the answer vary for the same question asked multiple times - Community

(2024) *OpenAI Developer Community*. Available at: https://community.openai.com/t/why-does-the-answer-vary-for-the-same-question-asked-multiple-times/770718 (Accessed: 27 February 2025).

World Economic Forum (2016) *Ten 21st-century Skills Every Student Needs*, *World Economic Forum*. Available at: https://www.weforum.org/agenda/2016/03/21st-century-skills-future-jobs-students/ (Accessed: 30 August 2023).

Zacher, H. *et al.* (2009) 'What do younger and older workers want to accomplish? Age-related differences in content and characteristics of occupational goals', *Zeitschrift fur personalpsychologie*, 8(4), pp. 191–200.

GLOSSARY OF TERMS AND DEFINITIONS

Actor

an entity capable of initiating, conducting, or withholding actions to influence its environment. Actors can be individuals or collective entities (such as organizations, institutions, states, or networks), operating at varying levels of complexity and agency. They interact within physical, a variety of social and/or ecological systems, shaping outcomes through decisions, behaviors, and strategic actions

Critical Thinking

Critical Thinking is the process of transforming data into information, analyzing information, making inferences, evaluating arguments, and solving problems through applying rigor logic, clear reasoning, intellectual dispositions, and ethical integrity. It involves recognizing and overcoming cognitive biases to develop a Mental Model that accurately reflects reality.

CT

Critical Thinking

Friendly Voice

a tone of communication that is perceived by the receiver as intented to be constructive and helping with the challenges the receiver is confromted with.

HEI

Higher Education Institute

Impact

the effect of an Actor's actions on the external phenomena deemed relevant within its Mental Model. It is inherently relative, shaped by the Actor's Self-concept, values, and worldview, which define the system's boundaries and determine what is considered meaningful change.

Mental Model

a conceptual representation of the Self and the reality around it, consisting of -highly simplified- an image of the "Self" and (un)conscious perceiving and feeling (LeDoux, 2020). Mental Models, above all, serve the "Self" to survive and thrive, and therefore to maintain a coherent perception of the Self and the world, even when it does not accurately reflect reality (Cline, 2016).

MM

Mental Model

MO

Manager / Other senior colleague

Model

an Actor-driven, purposefully simplified representation of reality that includes selected elements deemed important and the logic of their interactions. It serves to explain past phenomena, explore present dynamics, and estimate future outcomes. Models can be conceptual, mathematical, computational, or physical, and they function as tools for analysis, prediction, and decision-making by abstracting complexity while retaining relevant structure and behavior.

Self

the subjective experience of existence of an Actor, encompassing its awareness of identity, continuity, and agency. It includes a Self-concept—a cognitive and affective representation of one's attributes, roles, and relationships—shaped by internal states, interactions, and external influences. The Self enables an Actor to interpret its environment, make decisions, and engage in purposeful action.

Tone of voice

See Friendly Voice

YP

Young Professional

APPENDIX A

VARIABLES AND SURVEY QUESTIONS

VARIABLE (Question, answering possibilities and additional information)		**TRANSLATION** (with Deepl.com, manually corrected)
YPMatchesMO (Automated, but sometimes manually corrected)		
Het ID nummer van de MO die matcht met deze YP	Coding	The ID number of the MO matching this YP
Nummer	0-100	Number
YPTime_Minutes		
Invultijd, in minuten	Coding	The time it took to complete the form, in minutes
Numeriek	0-endless	Number
YPTeamSize		
Teamgrootte: met hoeveel collega's ongeveer werk je direct dagelijks samen?	Coding	Team size: approximately how many colleagues do you work with directly on a daily basis?
1 tot en met 5 collega's	1	1 to 5 colleagues
6 tot en met 10 collega's	2	6 to 10 colleagues
11 tot en met 20 collega's	3	11 to 20 colleagues
21 tot en met 50 collega's	4	21 to 50 colleagues
51 collega's of meer	5	51 colleagues or more
NietGevonden	Fout	NotFound
YPMan		
Heb je op dit moment een leidinggevende functie?	Coding	Are you currently in a managerial position?
Ja	Y	Yes
Nee	N	No
NietGevonden	Fout	NotFound
YPEduLevel		
(Source: YPEdu, coded by hand afterwards)		
Welke "hoogste" opleiding heb je succesvol afgerond?		What ‘highest’ education did you successfully complete?

(Bedoeld wordt een reguliere opleiding aan een door de Nederlandse overheid erkend instituut, zoals een Hogeschool of Wetenschappelijke Universiteit. Indien je meerdere opleidingen hebt afgerond, benoem de meest recente. Graag de opleidingsnaam voluit vermelden)	Coding	(Meaning a regular education at an institute recognised by the Dutch government, such as a UAS or Scientific University. If you have completed several courses, please mention the most recent one. Please state the name of the course in full)
HBO	HBO	University of Applied Sciences
WO	WO	Research University
Gepromoveerd (PHD/DBA/Anders)	PROM	PHD/DBA/Other
Anders	Other	Other
YPJobMonths		
Hoeveel maanden (inclusief eventuele proeftijd) ben je, gerekend vanaf het moment dat je wist dat je geslaagd was, inmiddels werkzaam?		Counting from the time you knew you had passed, how many months (including any probationary period) have you now been employed?
(NB: hier wordt een dienstbetrekking bedoeld die opleidings-relevant is. Dit mag ook een deeltijdbetrekking zijn, ook mogen periodes bij andere werkgevers meegerekend worden. Geef deze totale periode s.v.p. als een getal, in maanden, weer)	Coding	(NB: this refers to employment that is training-relevant. This may also be a part-time job or periods of employment with other employers. Please express this total period as a number, in months).
Numeriek	0-100	Number
YPEduProfit		
Waar heb je, met betrekking tot die "hoogst genoten" opleiding, tot op dit moment nog nadrukkelijk profijt van?		With regard to that ‘highest education’, where have you benefited emphatically up to this point?
YPEduProfitDipl		
Ik heb een diploma dat toegang geeft tot een baan op mijn niveau;	Coding	I have a diploma that gives access to a job at my level
Boolean	T/F	Boolean
YPEduProfitKnowl		
Ik heb daar vakkennis en beroepsvaardigheden opgedaan die ik nu nog dikwijls gebruik;	Coding	I gained professional knowledge and skills there that I still use often today
Boolean	T/F	Boolean

YPEduProfitThink		
Ik ben daar in het algemeen op een andere manier gaan denken en pas dat bewust en onbewust toe;	Coding	I started thinking in a different way there in general and apply that consciously and subconsciously
Boolean	T/F	Boolean
YPEduProfitNetwork		
Ik heb daar een netwerk van vrienden en kennissen opgedaan waar ik op kan terug vallen;	Coding	I gained a network of friends and acquaintances there that I can fall back on
Boolean	T/F	Boolean
YPEduProfitOther		
Anders	Coding	Other
Vrije tekst	Tekst	Free text
In hoeverre zijn de volgende stellingen naar jouw mening van toepassing?		In your opinion, to what extent do the following statements apply?
YPImpactEnvir		
Ik maak significante impact op mijn werk-omgeving	Coding	I make significant impact on my working environment
Helemaal oneens	1	Totally disagree
Vooral oneens	2	Mostly disagree
Neutraal	3	Neutral;
Vooral eens	4	Mostly agree
Helemaal eens	5	Totally agree
NietGevonden	Fout	NotFound
YPBelonging		
Ik voel mij echt volledig opgenomen in het team	Coding	I really feel fully included in the team
Helemaal oneens	1	Totally disagree
Vooral oneens	2	Mostly disagree
Neutraal	3	Neutral
Vooral eens	4	Mostly agree
Helemaal eens	5	Totally agree
NietGevonden	Fout	NotFound
YPCTAllowed (NB: reversed coded!)		
Een Young Professional mag hier niet kritisch op de status quo zijn	Coding	A Young Professional may not be critical of the status quo here

Helemaal oneens	5	Totally disagree
Vooral oneens	4	Mostly disagree
Neutraal	3	Neutral
Vooral eens	2	Mostly agree
Helemaal eens	1	Totally agree
NietGevonden	Fout	NotFound
YPCTFreedom		
Ik krijg hier de ruimte om een nieuwe impuls te geven	Coding	I am given the space here to reinvigorate
Helemaal oneens	1	Totally disagree
Vooral oneens	2	Mostly disagree
Neutraal	3	Neutral
Vooral eens	4	Mostly agree
Helemaal eens	5	Totally agree
NietGevonden	Fout	NotFound
YPReflectionMode		
Welke beschrijving geeft jouw werkwijze het best weer? Jij…	Coding	Which description best reflects your approach? You…
..neemt taken/opdrachten aan en voert ze conform instructies uit.	1	..Takes on tasks/orders and carries them out in accordance with instructions.
..vraagt bij een taak/opdracht door om zelf te begrijpen WAT er precies HOE WANNEER gedaan moet worden en herplant soms taken als dat handiger is.	2	..when given a task/task, asks deeper questions to understand for himself WHAT exactly HOW WHEN to do it and sometimes replans tasks if it is more convenient.
..vraagt door om zelf te begrijpen WAAROM volgens de opdrachtgever iets precies gedaan moet worden en kan soms afwijken van de werkwijze of de taak/opdracht als dat handiger is.	3	..asks deeper questions to understand for himself WHY exactly, according to the client, something needs to be done and may sometimes deviate from the working method or task/mission if it is more convenient.
..vraagt door om zelf de bredere samenhang van de uitdaging van de opdrachtgever te begrijpen, oppert alternatieven en zet daarbij geregeld ook anderen aan het denken.	4	..asks deeper questions to understand for himself the broader context of the client's challenge, suggests alternatives and regularly stimulates others in their thinking process.

..steunt anderen, door het stellen van de juiste vragen, met het vinden van samenhang in en duiden van hun eigen waarnemingen, gedachten en gevoelens.	5	..supports others, by asking the right questions, with finding coherence in and interpreting their own observations, thoughts and feelings.
NietGevonden	Fout	NotFound
YPChallengeScoreReported		
Welke score heb je voor de Redenatie Challenge gehaald?		What score did you achieve for the Reasoning Challenge?
Je kunt nu naar https://forms.office.com/e/Uwsj7wEzWJ om de redenatiechallenge te doen. Deze duurt maximaal 20 minuten. Vul na afloop svp het aantal behaalde punten hieronder in…	Coding	You can now go to https://forms.office.com/e/Uwsj7wEzWJ to do the reasoning challenge. It will take a maximum of 20 minutes. Afterwards, please fill in the number of points gained below....
Numeriek	0-100	Number
YPChallengeScoreGained (NB: using the real scoretables from the challenge-form itself)		
Welke score heb je voor de Redenatie Challenge gehaald?		What score did you achieve for the Reasoning Challenge?
Je kunt nu naar https://forms.office.com/e/Uwsj7wEzWJ om de redenatiechallenge te doen. Deze duurt maximaal 20 minuten. Vul na afloop svp het aantal behaalde punten hieronder in…	Coding	You can now go to https://forms.office.com/e/Uwsj7wEzWJ to do the reasoning challenge. It will take a maximum of 20 minutes. Afterwards, please fill in the number of points gained below....
Numeriek	0-100	Number
YPC10th		
Ben je geïnteresseerd in het (gratis) volgen van de "Red Teaming" cursus in februari/maart 2025, om daarmee de beschermde C10th titel te behalen? Er zijn maximaal 15 plekken beschikbaar.	Coding	Are you interested in attending the 'Red Teaming' course (for free) in February/March 2025, to obtain the protected C10th title? A maximum of 15 spots are available.
Ja, ik ben wel geïnteresseerd (en mijn eerder ingevulde emailadres	TRUE	Yes, I am interested (and my previously entered email address

mag gebruikt worden om mij hiervoor te benaderen)		may be used to contact me for this purpose)
Nee, dank, hier laat ik het bij…	FALSE	No thanks, I'll leave it at this…
DATAMO		
MOTime_Minutes		
Invultijd, in minuten	Coding	The time it took to complete the form, in minutes
Numeriek	0-endless	Number
MOMatchesYP (Automated, but sometimes manually corrected)		
Het ID nummer van de YP die matcht met deze manager/senior/overig persoon	Coding	The ID number of the YP matching this manager/senior/other person
Nummer	0-100	Number
MORelation		
Wat is uw werkrelatie ten opzichte van de Young Professional?	Coding	What is your working relationship towards the Young Professional?
Ik ben zijn/haar manager	1	I am his/her manager
Overig	2	Other
MOBranche	Coding	
Het antwoord zoals ingevoerd door de MO	Tekst	The answer, as given by the MO
MOISICS (Manual input)	Coding	
ISIC-codeomschrijvingen op niveau-1. Zie https://unstats.un.org/unsd/classifications/Econ/isic	AlfaNum	ISIC code descriptions at level-1. See https://unstats.un.org/unsd/classifications/Econ/isic
MOISICdesc	Coding	
Beschrijving van de branche (opgezocht nav MOISIC-veld)	Tekst	Industry description (looked up nav MOISIC field)
MOOrgSize		

Hoeveel medewerkers werken er ongeveer in totaal bij (en/of namens) uw organisatie?	Coding	Approximately how many employees in total work at (and/or on behalf of) your organisation?
Minder dan 10	1	Less than 10
10 of meer, maar minder dan 50	2	10 or more, but less than 50
50 of meer, maar minder dan 250	3	11 to 20 colleagues
250 of meer	4	250 or more
NietGevonden	Fout	NotFound
OMAnnTurnOver		
Hoeveel omzet, gegeven in euro's, draait uw organisatie ongeveer per jaar (of hoe groot is uw jaarbudget)?	Coding	Approximately how much turnover, given in euros, does your organisation turn over per year (or how big is your annual budget)?
2 miljoen of minder	1	2 million or less
Meer dan 2 miljoen, maar niet meer dan 10 miljoen	2	More than 2 million, but not more than 10 million
Meer dan 10 miljoen, maar niet meer dan 50 miljoen	3	More than 10 million, but not more than 50 million
Meer dan 50 miljoen	4	More than 50 million
NietGevonden	Fout	NotFound
MOOrgCat		
De grootte volgens de MKB definitie		Size according to the SME definition, Dutch chamber of commerce
Micro (minder dan 10 werknemers, minder dan 2 miljoen omzet)	1	Micro (fewer than 10 employees, less than 2 million turnover)
Klein (minder dan 50 werknemer, hoogstens 10 miljoen omzet)	2	Small (less than 50 employee, maximum turnover of 10 million)
Middelgroot (minder dan 250-werknemer, hoogstens 50 miljoen omzet)	3	Medium-sized (less than 250-employee, at most 50 million turnover)
Groot (250 of meer werknemers, Meer dan 50 miljoen omzet)	4	Large (250 or more employees, More than 50 million turnover)
In hoeverre zijn de volgende stellingen met betrekking tot de Young Professional (YP) naar uw mening van toepassing?		In your opinion, to what extent do the following statements regarding the Young Professional (YP) apply?
MOImpactEnvirYP		

De impact die deze YP op zijn/haar werk-omgeving maakt, is significant	Coding	The impact this YP makes on his/her working environment is significant
Helemaal oneens	1	Totally disagree
Vooral oneens	2	Mostly disagree
Neutraal	3	Neutra;
Vooral eens	4	Mostly agree
Helemaal eens	5	Totally agree
NietGevonden	Fout	NotFound
MOBelongingYP		
Deze YP is echt "1 van ons"	Coding	This YP is truly ‘1 of us’
Helemaal oneens	1	Totally disagree
Vooral oneens	2	Mostly disagree
Neutraal	3	Neutral
Vooral eens	4	Mostly agree
Helemaal eens	5	Totally agree
NietGevonden	Fout	NotFound
MOCTAllowedYP (NB: reversed coded!)		
Een YP mag hier niet kritisch op de status quo zijn	Coding	A YP may not be critical of the status quo here
Helemaal oneens	5	Totally disagree
Vooral oneens	4	Mostly disagree
Neutraal	3	Neutral
Vooral eens	2	Mostly agree
Helemaal eens	1	Totally agree
NietGevonden	Fout	NotFound
MOCTFreedomYP		
Een YP krijgt bij ons de ruimte om een nieuwe impuls te geven	Coding	A YP gets the space to reinvigorate with us
Helemaal oneens	1	Totally disagree
Vooral oneens	2	Mostly disagree
Neutraal	3	Neutral
Vooral eens	4	Mostly agree
Helemaal eens	5	Totally agree
NietGevonden	Fout	NotFound

MOReflectionModeYP		
Welke beschrijving geeft de werkwijze van de Young Professional naar uw mening het best weer? De YP…	Coding	In your opinion, which description best reflects the Young Professional's approach? The YP...
..neemt taken/opdrachten aan en voert ze conform instructies uit.	1	..Takes on tasks/orders and carries them out in accordance with instructions.
..vraagt bij een taak/opdracht door om zelf te begrijpen WAT er precies HOE WANNEER gedaan moet worden en herplant soms taken als dat handiger is.	2	..when given a task/task, asks deeper questions to understand for himself WHAT exactly HOW WHEN to do it and sometimes replans tasks if it is more convenient.
..vraagt door om zelf te begrijpen WAAROM volgens de opdrachtgever iets precies gedaan moet worden en kan soms afwijken van de werkwijze of de taak/opdracht als dat handiger is.	3	..asks deeper questions to understand for himself WHY exactly, according to the client, something needs to be done and may sometimes deviate from the working method or task/mission if it is more convenient.
..vraagt door om zelf de bredere samenhang van de uitdaging van de opdrachtgever te begrijpen, oppert alternatieven en zet daarbij geregeld ook anderen aan het denken.	4	..asks deeper questions to understand for himself the broader context of the client's challenge, suggests alternatives and regularly stimulates others in their thinking process.
..steunt anderen, door het stellen van de juiste vragen, met het vinden van samenhang in en duiden van hun eigen waarnemingen, gedachten en gevoelens.	5	..supports others, by asking the right questions, with finding coherence in and interpreting their own observations, thoughts and feelings.
NietGevonden	Fout	NotFound
MOCTBehave1YP		

Neem de volgende twee uitersten A en B: A) Accepteert klakkeloos wat anderen zeggen en B) Weegt het belang wat de informatiebron heeft mee, vraagt bewijs en controleert gegeven informatie. Welk gedrag neemt u bij de Young Professional voornamelijk waar?	Coding	Take the following two extremes A and B: A) Accepts blindly what others say and B) Weighs the importance what the information source has, asks for proof and checks given information. What behaviour do you observe in the Young Professional mainly?
A	1	A
<	2	<
-	3	-
>	4	>
B	5	B
NietGevonden	Fout	
MOCTBehave2YP		
Neem de volgende twee uitersten A en B: A) Zoekt alleen naar bewijs dat past bij zijn/haar mening en B) Zoekt actief naar bewijs dat de eigen mening weerspreekt en past zonodig zijn/haar oordeel aan Welk gedrag neemt u bij de Young Professional voornamelijk waar?	Coding	Take the following two extremes A and B: A) Looks only for evidence that fits his/her opinion and B) Actively looks for evidence that contradicts his/her own opinion and adjusts his/her judgement if necessary What behaviour do you observe in the Young Professional mainly?
A	1	A
<	2	<
-	3	-
>	4	>
B	5	B
NietGevonden	Fout	
MOCTBehave3YP		

Neem de volgende twee uitersten A en B: A) Stelt geen verdiepende vragen en B) Stelt positief-kritische vragen en vraagt door Welk gedrag neemt u bij de Young Professional voornamelijk waar?	Coding	Take the following two extremes A and B: A) Does not ask probing questions and B) Asks positive-critical questions and probing questions What behaviour do you observe in the Young Professional mainly?
A	1	A
<	2	<
-	3	-
>	4	>
B	5	B
NietGevonden	Fout	
MOCTBehave4YP		
Neem de volgende twee uitersten A en B: A) Trekt overhaaste conclusies, denkt zwart-wit en slaat complexe problemen (te) plat en B) Neemt de tijd om meerdere perspectieven te onderzoeken en bewijs te verzamelen voordat hij/zij tot een afgewogen conclusie komt Welk gedrag neemt u bij de Young Professional voornamelijk waar?	Coding	Take the following two extremes A and B: A) Draws hasty conclusions, thinks in black and white and flattens complex problems (too much) and B) Takes time to explore multiple perspectives and gather evidence before coming to a considered conclusion What behaviour do you observe in the Young Professional mainly?
A	1	A
<	2	<
-	3	-
>	4	>
B	5	B
NietGevonden	Fout	
MOCTBehave5YP		

Neem de volgende twee uitersten A en B: A) Trekt op basis van drogredenen en intern tegenstrijdige logica onlogische conclusies en B) Trekt logische conclusies op basis van duidelijke, samenhangende en rationele argumenten Welk gedrag neemt u bij de Young Professional voornamelijk waar?	Coding	Take the following two extremes A and B: A) Draws illogical conclusions based on fallacies and internally contradictory logic and B) Draws logical conclusions based on clear, coherent and rational arguments What behaviour do you observe in the Young Professional mainly?
A	1	A
<	2	<
-	3	-
>	4	>
B	5	B
NietGevonden	Fout	
MOCTBehave6YP		
Neem de volgende twee uitersten A en B: A) Trekt conclusies op basis van emoties en B) Redeneert op basis van feiten Welk gedrag neemt u bij de Young Professional voornamelijk waar?	Coding	Take the following two extremes A and B: A) Draws conclusions based on emotions and B) Reasons based on facts What behaviour do you observe in the Young Professional mainly?
A	1	A
<	2	<
-	3	-
>	4	>
B	5	B
NietGevonden	Fout	
MOCTBehave7YP		
Neem de volgende twee uitersten A en B: A) Wijst kritische feedback af, ook als deze positief bedoeld is en B) Ziet weloverwogen kritiek als kans tot verbeteren	Coding	Take the following two extremes A and B: A) Rejects critical feedback, even if it is meant positively and B) Sees considered criticism as an opportunity for improvement

Welk gedrag neemt u bij de Young Professional voornamelijk waar?		What behaviour do you observe in the Young Professional mainly?
A	1	A
<	2	<
-	3	-
>	4	>
B	5	B
NietGevonden	Fout	
MOCTBehave8YP		
Neem de volgende twee uitersten A en B: A) Volgt zonder zelf nadenken de overtuiging van de groep en B) Denkt zelfstandig en onafhankelijk, wijkt niet af voor groepsdruk Welk gedrag neemt u bij de Young Professional voornamelijk waar?	Coding	Take the following two extremes A and B: A) Follows the group's beliefs without thinking for him/herself and B) Thinks independently and independently, does not give in to group pressure What behaviour do you observe in the Young Professional mainly?
A	1	A
<	2	<
-	3	-
>	4	>
B	5	B
NietGevonden	Fout	
MOBehaveYPTotal		
Berekend veld, dat het waargenomen CT gedrag van de YP samenvat. Bestaat uit 4 keer de waarde van MOReflectionModeYP (in totaal dus max 5pt*4 = 20 punten) en alle MOBehave*YP bij elkaar (is dus max 5pt*8 vragen = 40 punten), gedeeld door het mogelijke punten (60), maal 100. Is dus een	Code	Calculated field, which summarises the observed CT behaviour of the YP. Consists of 4 times the value of MOReflctionModeYP (so in total, max 5pt*4 = 20 points) and all MOBehave*YP together (so is max 5pt*8 questions = 40 points), divided by the possible points (60), times 100. So is a number from 0 to 100.

getal van 0 tot 100. Afgerond op een geheel getal. Zo telt de reflectiemodeus voor in tottaal 1/3 mee. Dit is puur gevoelsmatig...		Rounded to an integer. This way the reflectionmodus weights for 1/3. This is purely based on "Feeling"
Nummer	0-100	Number
MOImpactOrg		
In hoeverre is de bijdrage van de Young Professional (YP) merkbaar (geweest) voor uw organisatie(onderdeel) zelf?		To what extent has the contribution of the Young Professional (YP) been (been) noticeable for your organisation/organisational unit itself?
MOImpactOrgNone		
Eigenlijk is in de periode dat de YP hier werkt (vrijwel) niets merkbaar veranderd voor mijn organisatie(onderdeel);	Coding	Actually, in the period the YP has been working here, (almost) nothing has noticeably changed for my organisation/organisational unit
Boolean	T/F	Boolean
MOImpactOrgFeeling		
Wij zijn ons bewuster geworden van onze "ongeschreven regels" en groepsdynamiek;	Coding	We have become more aware of our ‘unwritten rules’ and group dynamics
Boolean	T/F	Boolean
MOImpactOrgPerceiving		
Wij hebben onze eigen "Body of Knowledge" explicieter gemaakt en/of onze kennis uitgebreid;	Coding	We have made our own ‘Body of Knowledge’ more explicit and/or expanded our knowledge
Boolean	T/F	Boolean
MOImpactOrgSelf		
Ons idee van wie we zijn en wat we kunnen bijdragen aan de buitenwereld is veranderd;	Coding	Our idea of who we are and what we can contribute to the outside world has changed
Boolean	T/F	Boolean
MOImpactOrgOther		
Overig	Coding	Other
Tekst	Tekst	Tekst
MOImpactOrgTotal		

Berekend veld, dat de Impact van de YP op de organisatie samenvat. Bestaat uit 0 als er geen Impact wordt gerapporteerd. Anders voor ieder (perceiving, feeling, other) 1 punt, en het Self 2 punten. In totaal dus maximaal 6 punten. Dit is puur gevoelsmatig...	Code	Calculated field, summarising the Impact of the YP on the organisation. Consists of 0 if no Impact is reported. Otherwise for each (perceiving, feeling, other) 1 point, and the Self 2 points. So a total maximum of 6 points. This formula is purely based on "Feeling"
Nummer	0-5	Number
MOImpactEnvir		
In hoeverre is de bijdrage van de Young Professional (YP) merkbaar (geweest) voor de buitenwereld van uw organisatie(onderdeel)?		To what extent has the contribution of the Young Professional (YP) been (been) noticeable to the outside world of your organisation/organisational unit?
MOImpactEnvirNone		
Eigenlijk is er in de periode dat de YP hier werkt (vrijwel) niets voor de buitenwereld merkbaar veranderd;	Coding	Actually, during the period the YP has been working here, (almost) nothing has noticeably changed for the outside world
Boolean	T/F	Boolean
MOImpactEnvirH0		
Door de bijdrage van de YP is de wijze waarop onze producten/diensten tot stand komen gewijzigd (denk aan: doelmatiger, doeltreffender, veiliger, schoner, inclusiever, et cetera), waar de buitenwereld indirect voordeel bij heeft;	Coding	The YP's contribution has changed the way our products/services are created (E.g: more efficient, more effective, safer, cleaner, more inclusive, et cetera), which indirectly benefits the outside world
Boolean	T/F	Boolean
MOImpactEnvirH1		
Door de bijdrage van de YP zijn onze producten/diensten zelf, voor de buitenwereld merkbaar, veranderd;	Coding	As a result of the YP's contribution, our products/services have themselves, to the outside world, noticeably changed
Boolean	T/F	Boolean
MOImpactEnvirH2		

Door de bijdrage van de YP is onze reden van bestaan gewijzigd, wat gevolgen heeft voor het businessmodel en/of de samenwerkingsvormen en/of partijen met wie we samenwerken..;	Coding	The YP's contribution has changed our reason for existence, affecting the business model and/or the partnerships and/or parties we work with....
Boolean	T/F	Boolean
MOImpactEnvirOther		
Anders	Coding	Other
Vrije tekst	Tekst	Free text
MOImpactEnvirTotal		
Berekend veld, dat de Impact van de YP op de omgeving samenvat. Bestaat uit 0 als er geen Impact wordt gerapporteerd. Anders voor H0 (beter proces) 1 punt, H1 (product/dienst) 2 punten, H3 (nieuwe business) 3 punten, Other 1 punt. Dus een totaal maximaal van 6 punten. Dit is puur gevoelsmatig...	Code	Calculated field, summarising the Impact of the YP on the environment. Consists of 0 if no Impact is reported. Otherwise for H0 (better process) 1 point, H1 (product/service) 2 points, H3 (new business) 3 points, Other 1 point. So a total maximum of 6 points. This formula is purely based on "Feeling"
Nummer	0-5	Number
MOImpactFree		
Vrije tekst, zelf op te geven…	Coding	Other
Vrije tekst	Tekst	Free text
MOEduLevel		
Wat is uw eigen hoogst genoten opleiding?		What your own ‘highest’ education level?
(NB: de vraag is net anders dan die aan de YP, dus PHD kan ook onder "anders" vallen hier..)	Coding	(Note: the question differs from the onde asekd to the YP, so PHD may also be "Other" here…)
HBO	HBO	University of Applied Sciences
WO	WO	Research Univeristy
Anders	Other	Other
MOEduProfit		
Waar heeft u, met betrekking tot die "hoogst genoten" opleiding, ook op dit moment nog aanzienlijk profijt van?		With regard to that ‘highest education’, where have you benefited greatly up to this point?
MOEduProfitDipl		

Ik heb een diploma dat toegang geeft tot een baan op mijn niveau;	Coding	I have a diploma that gives access to a job at my level
Boolean	Y/N	Boolean
MOEduProfitKnowl		
Ik heb daar vakkennis en beroepsvaardigheden opgedaan die ik nu nog dikwijls gebruik;	Coding	I gained professional knowledge and skills there that I still use often today
Boolean	Y/N	Boolean
MOEduProfitThink		
Ik ben daar in het algemeen op een andere manier gaan denken en pas dat bewust en onbewust toe;	Coding	I started thinking in a different way there in general and apply that consciously and subconsciously
Boolean	Y/N	Boolean
MOEduProfitNetwork		
Ik heb daar een netwerk van vrienden en kennissen opgedaan waar ik op kan terug vallen;	Coding	I gained a network of friends and acquaintances there that I can fall back on
Boolean	Y/N	Boolean
MOEduProfitOther		
Anders	Coding	Other
Vrije tekst	Tekst	Free text
MOImpactTotal		
Is MOImpactOrgTotal plus 2 maal MOImactEnvirTotal. Deze formule is gevoelsmatig (Impact op de omgeving telt dubbel). Maximaal dus 5 plus 2*7 = 19 punten…		Is MOImpactOrgTotal and twice MoImpactEnvirTotal (because the latter is more important), This formula is based on "guts". Maximum 5 plus 2 times 7 = 19 points in total
Nummer	0-19	Number

APPENDIX B

EXAMPLE OF TRANSLATION MS FORMS OUTPUT

Microsoft Forms gives the possibility to download results. However, when downloaded the Column names are the answered questions in full, as are the answers. For example: ticking a radio button in question 1 “What is your favourite colour” out of three choices “Red”,”Blue”,”Purple” does not lead to the result “[[1],[A]]”, but gives the downloaded result “[[What is your favourite colour],[Red]]”. This makes calculation of the calculated fields, as well as statistical analyses impossible. So, this table must be “translated”.

Below an example of the “translation” of 1 row of a response of a Young Professional is shown. To keep this example readable, only the first row, as is downloaded from Microsoft Forms, is shown. Again, for readability, the row is transposed below (row became column, column became row). For the translation the information of Appendix A was used, placed in a separate sheet named “Legenda”.

ORIGINAL DOWNLOADED FORMAT

Variable	Value
ID	1
Begintijd	45618,5985416667
Tijd van voltooien	45618,6212384259
E-mail	anonymous
Naam	
Tijd van laatste wijziging	
Ik ga akkoord met het gebruik van deze gegevens voor wetenschappelijk onderzoek. (NB: een volledige beschrijving van wat uw akkoord inhoudt, kunt u vinden op https://edu.nl/fypvv).	Ja, ik ga akkoord met deze voorwaarden
De zoeksleutel:	<*nameremoved*>@gmail.com
Ik beantwoord deze vragenlijst in de rol van:	Young Professional (YP)
Teamgrootte: met hoeveel collega's ongeveer werk je direct dagelijks samen?	11 tot en met 20 collega's
Heb je op dit moment een leidinggevende functie?	Ja

Welke "hoogste" opleiding heb je succesvol afgerond?	WO Master of Science Leisure and Tourism Studies
Hoeveel maanden (inclusief eventuele proeftijd) ben je, gerekend vanaf het moment dat je wist dat je geslaagd was, inmiddels werkzaam?	36
Waar heb je, met betrekking tot die "hoogst genoten" opleiding, tot op dit moment nog nadrukkelijk profijt van?	Ik heb een diploma dat toegang geeft tot een baan op mijn niveau;Ik ben daar in het algemeen op een andere manier gaan denken en pas dat bewust en onbewust toe;Ik heb daar een netwerk van vrienden en kennissen opgedaan waar ik op kan terug vallen;Ik heb daar vakkennis en beroepsvaardigheden opgedaan die ik nu nog dikwijls gebruik;
Ik maak significante impact op mijn werk-omgeving	Vooral eens
Ik voel mij echt volledig opgenomen in het team	Helemaal eens
Een Young Professional mag hier niet kritisch op de status quo zijn	Vooral oneens
Ik krijg hier de ruimte om een nieuwe impuls te geven	Helemaal eens
Welke beschrijving geeft jouw werkwijze het best weer? Jij...	..vraagt door om zelf de bredere samenhang van de uitdaging van de opdrachtgever te begrijpen, oppert alternatieven en zet daarbij geregeld ook anderen aan het denken.
Welke score heb je voor de Redenatie Challenge gehaald?	84
Ben je geïnteresseerd in het (gratis) volgen van de "Red Teaming" cursus in februari/maart 2025, om daarmee de beschermde C10th titel te behalen? Er zijn maximaal 15 plekken beschikbaar.	Ja, ik ben wel geïnteresseerd (en mijn eerder ingevulde emailadres mag gebruikt worden om mij hiervoor te benaderen)

Heb je inmiddels aan een manager/senior collega een verzoek gestuurd om een indruk te geven van de impact die jij in je werkomgeving maakt?	Ja, ik heb inmiddels zo'n verzoek per mail verstuurd
Het onderzoek zal rond de zomer 2025 afgerond zijn. Zou je een samenvatting van de resultaten willen ontvangen?	Ja, ik ben wel geïnteresseerd (en mijn eerder ingevulde emailadres mag gebruikt worden om mij hiervoor te benaderen)

FORMULAS FOR TRANSLATION

Variable	Value
YPID	=[VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B25
YPTime_Minutes	=AFRONDEN(([VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B27-[VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B26)*24*60;0)
YPTeamSize	=VERT.ZOEKEN([VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B34;[VragenlijstVerbeteraarPer15jan2025.xlsx]Legenda!$A$9:$B$14;2;ONWAAR)
YPMan	=VERT.ZOEKEN([VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B35;[VragenlijstVerbeteraarPer15jan2025.xlsx]Legenda!$A$18:$B$20;2;ONWAAR)
YPEdu	=[VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B36
YPEduLevel	WO
YPJobMonths	=[VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B37
YPEduProfitDipl	=ISGETAL(VIND.SPEC("Diploma";[VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B38))
YPEduProfitKnowl	=ISGETAL(VIND.SPEC("vakkennis";[VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B38))
YPEduProfitThink	=ISGETAL(VIND.SPEC("denken";[VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B38))
YPEduProfitNetwork	=ISGETAL(VIND.SPEC("netwerk";[VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B38))
YPEduProfitOther	=SUBSTITUEREN(SUBSTITUEREN([VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B38;[VragenlijstVerbeteraarPer15jan2025.xlsx]Legenda!$A$39;"");[VragenlijstVerbeteraarPer15jan2025.xlsx]Legenda!$A$42;"") [VragenlijstVerbeteraarPer15jan2025.xlsx]Legenda!

YPImpactEnvir	=VERT.ZOEKEN([VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B39;[VragenlijstVerbeteraarPer15jan2025.xlsx]Legenda!$A$58:$B$63;2;ONWAAR)
YPBelonging	=VERT.ZOEKEN([VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B40;[VragenlijstVerbeteraarPer15jan2025.xlsx]Legenda!$A$66:$B$71;2;ONWAAR)
YPCTAllowed	=VERT.ZOEKEN([VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B41;[VragenlijstVerbeteraarPer15jan2025.xlsx]Legenda!$A$74:$B$79;2;ONWAAR)
YPCTFreedom	=VERT.ZOEKEN([VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B42;[VragenlijstVerbeteraarPer15jan2025.xlsx]Legenda!$A$82:$B$87;2;ONWAAR)
YPReflectionMode	=VERT.ZOEKEN([VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B43;[VragenlijstVerbeteraarPer15jan2025.xlsx]Legenda!$A$91:$B$96;2;ONWAAR)
YPChallengeScoreReported	=[VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B44
YPChallengeScoreGained	
YPC10th	=ISGETAL(VIND.SPEC("Ja";[VragenlijstVerbeteraarPer15jan2025.xlsx]GegYP!B45))

OUTPUT, READY FOR ENRICHMENT AND STATISTICAL ANALYZIS

Variable	Value
YPID	1
YPTime_Minutes	33
YPTeamSize	3
YPMan	Y
YPEdu	WO Master of Science Leisure and Tourism Studies
YPEduLevel	WO
YPJobMonths	36
YPEduProfitDipl	WAAR
YPEduProfitKnowl	WAAR
YPEduProfitThink	WAAR
YPEduProfitNetwork	WAAR
YPEduProfitOther	
YPImpactEnvir	4
YPBelonging	5

YPCTAllowed	4
YPCTFreedom	5
YPReflectionMode	4
YPChallengeScoreReported	84
YPChallengeScoreGained	
YPC10th	WAAR

NB: further steps (Manual enrichment, calculation of the calculated variables) are not shown here.

APPENDIX C

CT ABILITY TEST: CORRECT ANSWERS

Question	Password	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
2		C	B	A	B	C	C	C	A	C	C	A	B	B	A	B	A	A	A	C	B	C	C	D	A	A	A
3		D	A	B	B	A	C	D	A	C	B	D	D	C	B	A	D	B	A	B	C	D	C	C	A	B	D
4		C	B	A	D	B	D	B	C	A	C	B	A	C	C	A	C	D	B	B	B	D	A	D	A	A	A
5		C	C	B	D	D	A	D	A	C	B	D	B	A	A	B	B	B	C	B	A	D	D	D	C	B	A
6		D	A	D	C	B	B	D	B	D	D	A	C	C	C	C	B	C	C	D	D	B	C	B	C	D	D
7		D	D	A	B	D	C	A	B	C	B	B	B	A	D	A	B	D	B	A	B	A	C	A	C	D	D
8		C	D	C	C	B	B	D	B	B	D	A	B	C	D	D	C	B	C	B	A	D	D	D	D	A	C
9		C	A	C	D	A	B	A	A	C	A	C	C	A	C	C	D	C	D	C	C	D	D	D	D	B	D
10		D	B	D	D	B	D	A	D	D	C	C	D	A	D	C	D	A	B	A	C	B	B	D	B	B	D
11		D	C	B	B	D	C	A	D	B	C	C	C	C	D	B	C	B	D	A	D	A	A	B	B	A	B
12		C	B	D	D	D	A	A	B	B	A	C	B	C	A	B	A	D	D	D	D	D	B	B	B	B	D
13		B	D	D	D	C	B	A	B	C	D	A	D	B	B	D	C	D	D	C	B	A	B	A	B	B	D
14		A	A	D	C	D	C	D	D	D	B	C	C	B	D	C	C	D	B	C	C	D	C	D	C	B	A
15		B	D	D	A	C	B	B	D	D	C	B	D	D	D	D	A	C	D	A	C	C	B	B	D	A	D
16		D	B	B	C	B	A	A	B	D	A	D	C	C	D	A	D	A	A	C	A	C	D	C	D	C	D
17		B	A	B	B	A	D	B	A	C	A	C	A	A	A	C	B	A	C	C	D	B	D	D	A	A	B
18		B	D	A	C	C	D	A	D	B	C	B	A	C	B	B	B	C	C	B	B	D	A	A	C	B	C
19		B	C	A	D	C	A	C	B	B	C	D	A	D	D	B	B	A	C	B	B	D	D	A	D	A	C
20		A	D	A	B	C	B	C	C	B	B	D	D	B	A	A	D	C	B	C	B	A	A	D	C	B	B
21		A	C	B	D	C	D	D	D	C	B	B	D	D	B	D	A	A	D	B	A	D	D	A	A	C	B
22		A	B	B	B	A	D	C	A	C	B	B	C	B	A	B	C	C	B	D	A	D	C	A	B	A	D
23		D	A	A	D	C	D	B	A	D	B	D	A	B	C	A	D	C	A	A	D	D	D	C	C	A	B
24		A	C	B	D	A	D	C	A	D	C	D	C	A	A	C	C	D	B	A	C	D	D	A	C	C	D
25		A	C	B	B	B	D	B	B	B	A	D	B	C	C	C	C	B	B	B	D	A	B	D	D	C	C
26		B	A	A	D	C	A	D	B	D	C	C	C	A	C	A	C	C	A	C	A	D	A	A	B	B	C

If you want to check the answer question by question: fill in the password from top to bottom, and the right answer can be found at the intersection of the question number (the row) and the password letter (the column)

However, if you take the paper version of the ability test all at once and don't check your answers question by question, the right answers are: 2C; 3D; 4B; 5A; 6C; 7A; 8D; 9A; 10D; 11C; 12A; 13D; 14D; 15A; 16A; 17B; 18A; 19B; 20A; 21A; 22B; 23A; 24A; 25A; 26A

APPENDIX D

CONSENT FORM

Data was collected via MS-Forms. Every respondent agreed to the Consent Form, that was made available online via https://edu.nl/fypvv. In this form the participation, rights and procedures were communicated in both English and Dutch. In order to recognize the text of the form in this report, the original type font is not altered. The English text was as follows:

Participant Information Sheet/Consent Form

Social Science Research - *Adult providing own consent*

Title	Critical Thinking: prerequisite for making Impact?
Short Title	C10th
Project Sponsor	N.A.
Coordinating Principal Investigator/ Principal Investigator	drs. ing. B. de Graaf CIPM / CIPP/e
Associate Investigator(s)	N.A.
Location	Rotterdam

Let op: een geautomatiseerde Nederlandse vertaling is als bijlage bijgevoegd.

Part 1 What does my participation involve?

1 Introduction

You are invited to take part in this research project, which is called C10th. You have been invited because you are a recently Dutch higher education graduated Young Professional (YP). Your contact details were obtained by asking your email-address at the graduation ceremony you visited.

.
This Participant Information Sheet/Consent Form tells you about the research project. It explains the processes involved with taking part. Knowing what is involved will help you decide if you want to take part in the research.

Please read this information carefully. Ask questions about anything that you don't understand or want to know more about. Before deciding whether or not to take part, you might want to talk about it with a relative, friend or colleague.

Participation in this research is voluntary. If you don't wish to take part, you don't have to.

If you decide you want to take part in the research project, you will be asked to sign the consent section OR to tick the box "Ik neem vrijwillig en bewust deel aan dit onderzoek en ga er mee akkoord dat de gegevens geanonimiseerd worden gebruikt voor wetenschappelijk onderzoek." on the online questionnaire form. By signing it you are telling us that you:
• Understand what you have read
• Consent to take part in the research project
• Consent to be involved in the research described
• Consent to the use of your personal and health information as described.

You will be given a copy of this Participant Information and Consent Form to keep OR you may download it using the link on the online questionnaire form.

2 What is the purpose of this research?

The results of this research will be used by the researcher Boudewijn de Graaf to obtain a Doctoral Business Administration degree, who initiated the research and also conduct the research. The research is partly funded out of the training budget, made available by the Business IT & Management department of the Rotterdam University of Applied Sciences (RUAS) and partly of the researcher's own budget. Main goal is to find out if Critical Thinking indeed, as presumed widely, is an important prerequisite to make Impact, especially for Young Professionals.

3 What does participation in this research involve?

Participation in this research project involves filling in an online questionnaire by the YP and his/her manager. The duration of this phase will be around four weeks, starting at the end of 2024. The questionnaire used for the YP slightly differs from the manager questionnaire. The results of questionnaire filled in by the YP and his/her manager will not be mutually interchanged by the researcher. If both questionnaires are received in good order and the YP wants to participate in an extra training, 15 YP-participants will be drawn from the pool of respondents. Only these 15 participants will get a short training to stimulate them to conduct critical thinking behaviour in a business context. And only these 15 participants (and their managers) will be asked to fill in the questionnaire again. The duration of the training phase will be approximately 2 months. The duration of the second questionnaire phase will be around four weeks.

All people that react will be given the opportunity to be informed about the results of this research at the end of the research project. There are no costs involved for respondents. More details about the planning and the research project can be requested from the researcher by sending an email to Boudewijn@kronenbeeck.nl.

Because you are a higher educated Young Professional you are eligible to take part in this project. Completing the questionnaire will take approximately 45 minutes for the YP, and 20 minutes for the Manager. This research project has been designed to make sure the researchers interpret the results in a fair and appropriate way and avoids jumping to conclusions. There are no costs associated with participating in this research project, nor will you be paid. However, participants in the training phase will get a free meal.

4 Other relevant information about the research project

Research will be conducted amongst the Young Professionals that graduated at the Rotterdam Business School in the summer of 2024. The first questionnaire round will serve as a "control group" for the 15 participants in the training phase. After completing this research project no follow up research is foreseen at the moment. If the results are interesting enough, the project may result in a publication. Also, if the results show that training enhances impact, the training may be continued and offered to other YP's. The certificate that the 15 participants receive (when they complete the training successfully) gibes them the right to use the "title" that comes along with it (C10th). This trademark

has been registered by the European Trademark Office (EUIPO). The ownership of this trademark however is not transferred to the participants.

The (anonymised) data may be transferred to the research department of the RUAS, and then follows their research data regime.

5 Do I have to take part in this research project?

Participation in any research project is voluntary. If you do not wish to take part, you do not have to. If you decide to take part and later change your mind, you are free to withdraw from the project at any stage.

If you decide you want to take part in the research project, you will be asked to sign the consent section OR to tick the box “Ik neem vrijwillig en bewust deel aan dit onderzoek en ga er mee akkoord dat de gegevens geanonimiseerd worden gebruikt voor wetenschappelijk onderzoek.” on the online questionnaire form. By signing it you are telling us that you:

- Understand what you have read
- Consent to take part in the research project
- Consent to be involved in the research described
- Consent to the use of your personal information as described.

You will be given a copy of this Participant Information and Consent Form to keep OR you may download it using the link on the online questionnaire form. If you do decide to take part, you will be given this Participant Information and Consent Form to sign and you will be given a copy to keep.

Your decision whether to take part or not to take part, or to take part and then withdraw, will not affect the relationship with the researcher or the RUAS.

6 What are the possible benefits of taking part?

We cannot guarantee or promise that you will receive any benefits from this research; however, possible benefits may include receiving a training and use of the “C10th” tile after successful completement of the training.

7 What are the possible risks and disadvantages of taking part?

Filling in the questionnaire might make you, or your manager, think about your impact. This might be a good starting point for a good conversation. However, the researcher will maintain your answers private and confidential.

When drawing (randomly) from the respondents to get participants for the training, three groups are formed. The researcher will form the groups in a way that the risks of exchanging competition sensitive information of their organization is minimized. When

trained, you may experience some embarrassment or other social stress. However, the trainer will maximize the “safe space” during the training.

Apart from this, no risks are foreseen.

8 What if I withdraw from this research project?

If you decide to leave the research project, the researchers will not collect additional personal information from you, although personal information already collected will be retained to ensure that the results of the research project can be measured properly and to comply with law. You should be aware that data collected up to the time you withdraw will form part of the research project results. If you do not want your data to be included, you must tell the researchers when you withdraw from the research project.

9 Could this research project be stopped unexpectedly?

This research project may be stopped unexpectedly for a variety of reasons. These may include reasons such as illness.

10 What happens when the research project ends?

This project will lead to a dissertation. Depending on the results of the project, a publication will be written. If the training adds to the Impacts Young Professionals make in their first job, a training program may be offered to others by the researcher. Participants may ask a (free) copy of the dissertation. The research data may (anonymized) be handed over to the research department of the RUAS (see paragraph 4).

Part 2 How is the research project being conducted?

11 What will happen to information about me?

By signing the consent form you consent to the research team collecting and using personal information about you for the research project. Any information obtained in connection with this research project that can identify you will remain confidential and stored encrypted on a computer. Your information will only be used for the purpose of this research project and it will only be disclosed with your permission, except as required by law. Anonimised data may be transferred to the RUAS. The anonymized data may also be analysed using AI. The use of AI may involve the data being processed outside the EU, however, due to the contract with the AI provider the data will not be used to train the AI model.

The personal information that the research team collects and uses is the information from the questionnaires, and the findings from the training sessions. If the results of this

research project will be published and/or presented, information will be provided in such a way that you cannot be identified, except with your express permission.

In accordance with the GDPR, you have the right to request access to the information about you that is collected and stored by the research team. You also have the right to request that any information with which you disagree be corrected. Please inform the research team member named at the end of this document if you would like to access your information.

12 Complaints and compensation

If you suffer any distress or psychological injury as a result of this research project, you should contact the research team as soon as possible. You will be assisted with arranging appropriate treatment and support.

The information you exchange during the training sessions is up to you: the researchers do not take any liability for negative (business) impact stemming from an information directly exchanged between training participants, nor will they take (financial) credits for the positive impact stemming from information exchange between participants.

13 Who is organising and funding the research?

This research project is being conducted by Boudewijn de Graaf, who initiated the research and conducts the research. The research is partly funded out of the training budget, made available by the Business IT & Management department of the Rotterdam University of Applied Sciences (RUAS) and partly of the researcher's own budget.

The researcher may benefit financially from this research project if, for example, the project leads to, or may be incorporated in, any commercial enterprise. You will not benefit financially from your involvement in this research project even if, for example, knowledge acquired from your information proves to be of commercial value. However, this only applies to information regarding Critical thinking and making Impact. Business relevant information exchanged by participants in training sessions will NOT be used for commercial purposed by the researchers.

The researcher will not receive a personal (financial) benefit from your direct involvement in this project.

14 Who has reviewed the research project?

The ethical aspects of this research project have been approved by the Supervisor/Mentor of *SSBM Geneva*.
This statement has been developed to protect the interests of people who agree to participate in human research studies.

15 Further information and who to contact

If you want any further information concerning this project or if you have any problems which may be related to your involvement in the project, you can contact the researcher on Boudewijn@kronenbeeck.nl

If you have any complaints about any aspect of the project, the way it is being conducted or any questions about being a research participant in general, then you may contact the supervisor of the project: velimir@velimirsrica.com, see also https://www.ssbm.ch/personnel/velimir-srica-phd/

Declaration of Consent by Participant

I have read the Participant Information Sheet (or I've used Deepl.com for a translation in a language that I understand; the translation in Dutch is added as an attachment). I understand the purposes, procedures and risks of the research described in the project. I have had an opportunity to ask questions and I am satisfied with the answers I have received. I freely agree to participate in this research project as described and understand that I am free to withdraw at any time during the project without affecting my future care. I understand that I may download a copy of this document.

I've given my approval by ticking the box on the research form. I also may sign this document and send a digital copy to Boudewijn@kronenbeeck.nl

Name of Participant (please scan) _____

Signature _____ Date _____

Declaration by Researcher†

I have given the opportunity to a further explanation of the research project, its procedures and risks and I believe that the participant has understood that explanation.

Boudewijn de Graaf
9 September 2024