From creativity to innovation

From creativity to innovation:
Understanding and improving the evaluation and
selection of ideas in educational settings

Published by ROA
Postbus 616
6200 MD Maastricht

ISBN: 978-90-5321-596-8

Printed in the Netherlands by Canon

From creativity to innovation: Understanding and improving the evaluation and selection of ideas in educational settings

DISSERTATION

to obtain the degree of Doctor at Maastricht University,
on the authority of the Rector Magnificus,
Prof. dr. Rianne M. Letschert
in accordance with the decision of the Board of Deans,
to be defended in public on
1 July 2021, at 10:00 hours

by

Kim Christine Maria van Broekhoven

Supervisors
Prof. Dr. Rolf van der Velden
Prof. Dr. Lex Borghans

Co-supervisor:
Dr. Barbara Belfi

Assessment Committee
Prof. Dr. Trudie Schils (Chair)
Prof. Dr. Andries de Grip
Prof. Dr. Onne Janssen, University of Groningen
Dr. Simone Ritter, Radboud University

Acknowledgements

Over the last four years, many great things happened in my life for which I am grateful. I explored long-dreamed hobbies from Salsa to Muay Thai, I started travelling abroad by myself, I welcomed my little nephew Kars, I lived in Australia and finished my PhD. To finish a PhD during a worldwide pandemic does not only require an optimistic attitude, it also requires a good connection with colleagues, family and friends. I could not have written this dissertation without you, and I would like to thank you here.

First, I would like to express my gratitude to the members of my assessment committee, Prof. Dr. Trudie Schils, Prof. Dr. Andries de Grip, Prof. Dr. Onne Janssen and Dr. Simone Ritter for their time and effort to review and evaluate my dissertation. Trudie, I would like to thank you for acting as the chairwoman of my committee, and for all your valuable advice throughout my PhD. Furthermore, I would also like to thank committee members Prof. Dr. Wim van de Donk, Prof. Dr. Ute Hülsheger, Prof. Dr. Evelyn Kroesbergen, and Dr. Arie van der Lugt for their time invested in reading this dissertation, and their questions.

Next, I would like to thank my supervisors Prof. Dr. Rolf van der Velden, Prof. Dr. Lex Borghans, and Dr. Barbara Belfi for their endlessly generous academic tips, advice, and feedback. Rolf, I am grateful for all the opportunities and responsibilities you have given me since I started at ROA. Throughout my journey at ROA, your trust and confidence in me empowered me to go beyond. Lex, I really appreciate the countless creative brainstorm sessions to tackle research challenges (not 'problems' but 'challenges' in your words). Your positive and flexible research attitude have taught me that research is a fun and creative quest to knowledge. Barbara, thank you for your professional and personal guidance throughout my PhD journey. Your positive attitude gave me the confidence to explore new topics in my research. I really enjoyed our (almost) weekly meetings to discuss all kind of things.

Then there are, of course, all the other great colleagues at ROA that I would like to thank for making me feel part of a ROA family that is always welcoming and inspiring. These lines are not sufficient to thank each and every one of you – I take the liberty to mention a few. Sander, from the start of my journey at ROA, you have taken me under your wings to teach me programming in STATA. I do not know how many times I walked into your office with STATA questions, picking your critical mind, or just for a chat. Thank you for sharing your knowledge with me and your endless support. Joyce, I remember when Rolf told me that I have to meet Joyce! Since we met, you have taught me countless things from valuable advice about teaching to personal insights from which I learned a lot, thank you. I want to give Esther, Mariëlle, Margo, Willeke H., Willeke K., Fleur, Silvana, Marion, Silvia, Miranda and Melissa a special shout-out for all their help in making my PhD journey a bit easier. A special thanks to Esther for designing this book, as well as for all your help in my PhD projects. I could not have collected data among 400 primary school children without your help. Former roomies, I would like to thank you for a fun and supportive office climate, especially in times of 'home office', I appreciate you even more: Alexander, Kars, Lynn, Marloes, Maria F., Maria P., Sanne, Tobias, and Victor.

Finally, thanks to all PhD students with whom I got to share great conversations and memories: Corrie, Dinah, Henrik, Johannes, Juan, Laura, Mariana, Martijn, Max, Olga, Stan, Steffi. During my PhD, I had the great opportunity to be part of an interfaculty collaboration to set up creativity projects within various faculties. A special thanks to Arie, Barbara, Eline, Roy and Scheltus. I truly enjoyed tutoring Creative Problem Solving together. Last, but not least, I would like to thank my former colleagues, Chanvi, David, Michelle, Rebecca and Sue for welcoming me to the University of South Australia and all their support when I was 'stuck' in Adelaide.

Next to all amazing people I have met in my PhD journey, I would also like to take a moment to thank family and friends that played an important role in supporting me over the last years. Mom, Dad, Danique, Edo and Martijn, thank you for your genuine interest and support. The world of academia might feel alien to you, but your questions and curiosity made me feel blessed. For example, when mom consulted the internet to look up what a PhD or Postdoc actually entails. A special thanks to mom and dad for your unconditional love. You have always given me the freedom to make my own decisions, and supported those decisions. I cannot thank you enough for this, because it gave me the confidence to listen to my 'gut' feeling and to create the life that I want. Danique, my dear sister, since we were little, I have always looked upon you as an amazing kind person. When I was dealing with a personal or professional issue, you would immediately put down what you were working on to help me out. Thank you for this. Edo, dear brother, I am grateful that you temporarily took me into your house to live with you and Kars. It gave us the opportunity to get closer to each other, and I got to know my little nephew better. Ria and Chris, my godmother and godfather, I feel blessed being your godchild. You always make me feel welcome in your house, I am thankful for our close relationship. I also would like to give a special thanks to my uncle Cees, the Professor in the family; I always valued your insights and advice about the strange world of academia, and you are not alone anymore.

I could not have completed my PhD without my dear friends. The pomodoro group initially started out as a peer pressure-writing group, but quickly turned into a support group for each other. Sanne, Inge, Maja, Raginee, Charlotte and Julia, you know that a PhD is - next to a professional journey – a personal journey that takes you into an emotional rollercoaster with ups and downs. You always kept food, walks and good company coming which made my PhD journey amazing. Thank you for riding this rollercoaster with me, and I hope to always stay in contact wherever the world takes us. My dear friends, Maud, Pety, Fenne, Lotte, Sandy, Linda, Roxanne, and Michelle, we are all so different but at the same time so alike. I cannot thank you enough for all our good conversations and distraction during countless dinners, festivals, and weekends away. I have always lived somewhere else, but as you know, I will always find my way back to you lovely ladies! I also feel blessed with my dear high school friends, Chloé, Maartje, Laura and Stephanie. We all went different ways after high school, but we still see each other. It does not matter if we have not seen each other for months, it always feels like how it was in high school. Thank you to Dre, Julia, Marlon, Patrick, Randy, Rian and Suzanne for countless weekends away for an Honours reunion. Whether it is in Dublin or in Madrid, we always have good laughs and deep conversations about life, with the occasionally political incorrect comment of Randy. Dear Anne, my dear friend from Sociology, both being proud 'Brabander' makes us inseparable. We always kept in contact even though we lived miles apart. Thank you for our beautiful friendship. Undoubtedly, I did not express my gratitude to everyone that provided valuable advice or a listening ear; thank you to all other people I met along my PhD journey and did not mention here.

Contents

1

Introduction

1.1 Motivation

Worldwide, policy makers and business leaders have recognized that a new way of thinking should be cultivated in education to develop solutions for complex problems, such as global warming or ageing of the population (IBM, 2010; OECD, 2019; World Economic Forum, 2016; 2018; 2020). Recently, the importance of creativity has been highlighted even more by health related changes caused by COVID-19, like the sudden need for more IC beds in hospitals due to a rise in patients. Creativity gives students the ability to see things from new perspectives, generate novel and useful ideas, raise a variety of questions, and come up with solutions to complex problems (Puccio, 2017; Sternberg & Lubart, 1999). As such, creativity is seen as a crucial competence in all levels of education, ranging from elementary to tertiary education (e.g., Matraeva, Rybakova, Vinichenko, Oseev, & Ljapunova, 2020; OECD, 2019).

However, one of biggest gaps in literature is that widely accepted definitions of creativity tend to focus solely on idea generation, and tend to ignore other facets of creative problem solving, such as idea evaluation and idea selection (see Figure 1.1). For example, Plucker, Beghetto and Dow (2004, p. 90) defined creativity as *the interaction among aptitude, process, and environment by which an individual or group produces a perceptible product that is both novel (original, new) and useful (appropriate, feasible) as defined within a social context*. However, generating or producing creative ideas seldom is the final goal of a creative endeavor. Rather, to successfully solve problems, it requires one or a few good ideas that really work, and work better than previous ones (e.g., Rietzschel, Nijstad, & Stroebe, 2006). This requires that people are able to evaluate their own and others ideas, choose ideas to develop further, and abandon those that are unlikely to be successful. Thus, being creative does not stop with idea generation: creative ideas must be *recognized* and *selected* (i.e., idea evaluation and idea selection).

During the phase of *idea evaluation*, ideas, concepts or solutions that are generated in the phase of idea generation are evaluated on one or more dimensions (e.g., feasibility and originality). To evaluate the ideas, they are compared to some standard/benchmark or to each other (e.g., Mumford, Lonergan, & Scott, 2002). Benedek et al., (2016) found that people tended to underestimate the creativity of ideas. Specifically, although people recognized the novelty of highly creative ideas, they tended to underestimate the appropriateness or feasibility of these ideas. As such, there are gaps in our knowledge whether people are able to recognize creative ideas, and whether creative ideas are valued or appreciated. Next, in *idea selection*, one or several creative ideas out of a larger idea pool are selected for further development or implementation (e.g., Faure, 2004). People tend to reject highly original or risky ideas, and are more likely to select ideas that are consistent with social norms, and easy to understand (e.g., Blair & Mumford, 2007; Rietzschel, Nijstad, & Stroebe, 2010; Van Damme, Anseel, Duyck, & Rietzschel, 2019). Thus, existing research suggests that both idea evaluation and idea selection are difficult parts of the creative problem solving process, but important, since they serve as the bridge that links the generation (i.e., creativity) and implementation (i.e., innovation) of creative ideas (Mumford, Mobley, Reiter-Palmon, Uhlman, & Doares, 1991). As such, errors in the evaluation and selection of creative ideas are common occurrences that can form a bottleneck of innovation.

In addition, there are numerous examples of creative ideas that were initially rejected or ignored. One example is a one-hundred year old solution for a sun-starved Norwegian town. The small town of Rjukan is deeply embedded in a narrow valley. The mountains that surround it block the sunlight from September to March, causing Rjukan residents to live in a permanent shadow for more than six months a year. More than a century ago, the local newspaper published a novel and radical idea to use a mirror to reflect sunlight onto the town. However, this idea never came to fruition and instead they went for a more mundane idea and built a cable car to transport residents to the top of the mountain for sunlight exposure. After nearly hundred years, in 2013, the radical mirror idea has been put in practice with three high-tech mirrors of 50 square meters on a ridge on Gaustatoppen mountain, brightening up the previously gloomy town center (Jordan, 2013, November 1). Another example is the initially rejected idea of FedEx. Yale student Fred Smith identified that the passenger route systems used for shipping were not efficient for making urgent shipments (e.g., medical supplies). He, therefore, proposed a system designed for faster shipment delivery by carrying packages at night when airports are not congested. However, this idea did not receive any positive response from the college professors who thought the idea was not feasible (Haddad & Roman, 2004). Fred Smith kept on working on this idea and figured out a way to be able to get packages delivered within a span of two days. As such, he turned this seemingly unfeasible into the successful company known as "Federal Express" (i.e., FedEx). These examples were at the boundary of feasibility in their time, and those kind of ideas forces technology to move on.

Of course, these two ideas are most likely just the tip of the iceberg: nobody knows how many truly creative and promising ideas were never implemented. Systematic research is needed to understand when people do or do not recognize and select creative ideas. The aim of this thesis is to gain a deeper understanding of and improve the process of idea evaluation and idea selection.[1]

1.2 What is creativity?

In creativity research, an idea is judged to be "creative" or "of high quality" if it is both potential original and feasible (Corazza, 2016; Runco & Jaeger, 2012). The lack of either of these two qualities makes the idea either mundane or irrelevant, and no longer creative. Thus, the creative process involves a search for ideas or solutions that fulfill both criteria. While most creativity research uses the dual criteria of *originality* and *feasibility* to identify "creativity", Litchfield, Gilson, and Gilson (2015) argue that this may not do justice to the complexity of the matter. In their analysis of what it means for an idea to be "creative", they address some of the different definitions that have been used for both originality and feasibility.

With regard to *originality* (or novelty), they note that, although "[I]deas are considered to be novel to the extent that they are uncommon in terms of either their task or social context" (p. 242). Several conceptions of novelty exist: (a) *newness* (within a particular context or domain), (b) *frequency* (or rather infrequency), and (c) *distance* (the degree to which an idea is different from current practice). Litchfield et al. (2015) noted that these three facets of novelty are not necessarily aligned: "For example, the idea to provide decorated facial tissue boxes for different seasons might not constitute a novel idea in terms of distance, even if it is entirely new to a firm or rarely mentioned" (pp. 242-243).

A similar distinction is made with regard to *feasibility*, which Litchfield et al. argue is only one facet of the broader dimension "usefulness". While feasibility refers to the ease with which an idea could be implemented, there is also the *effectiveness* of an idea, which refers to the extent to which an idea actually solves the problem (i.e., expected value or success of the idea). If people fail to recognize creative ideas, this is not only caused by a perceived lack of feasibility, but might also by caused by a perceived lack of effectiveness (Rietzschel et al., 2010). For example, the publishers who rejected the first Harry Potter book or restaurant owners who rejected the Kentucky Fried Chicken recipe believed that it was easy to implement, but did not believe that it was possible to gain profit from these ideas (Nemeth, 1997).

1 It can be argued that the evaluation and selection of raw ideas first occur at the individual level. As such, this thesis focuses on individual person evaluation and selection.

1.3 The tension between originality and feasibility

Most studies on idea evaluation and idea selection use some combination of originality and feasibility to evaluate and select ideas on their creativity. One reason why it may be difficult to identify or select ideas that score high on both originality and feasibility is that, in people's mind, the two are often perceived as being negatively correlated.

Several authors have indeed suggested that these two criteria may be seen as incompatible and represent a fundamental tension or paradox (Frederiksen & Knudsen, 2017; Miron-Spektor & Beenen, 2015; Miron-Spektor, Gino, & Argote, 2011; Mueller, Melwani, & Goncalo, 2012; Nijstad, De Dreu, Rietzschel, & Baas, 2010; Zacher, Robinson, & Rosing, 2016). For instance, in a meta-analysis of 20 studies, Nijstad, De Dreu, Rietzschel and Baas (2010) found that the average correlation between originality and feasibility was moderately negative (r = -.42). Although some studies have found positive correlations between originality and feasibility (e.g., Kohn, Paulus, Choi, 2010), and it is certainly possible for ideas to be both highly original and highly feasible, most ideas are not. Moreover, people often *perceive* a negative correlation or even an incompatibility between the two concepts, and tend to focus on one of the criteria for creativity only at the expense of the other (see Rietschel, Nijstad, & Stroebe, 2010; Runco & Charles, 1993).

There are several reasons for this perceived negative correlation between originality and feasibility. Many ideas that people come up with are simply minor adaptions of existing practices, and as such are usually high in feasibility and low in originality (Ward, 1994). Research has further found that the generation of original versus feasible ideas may be triggered by different kinds of tasks (Kapoor & Khan, 2019; Runco, Illies, & Eisenman, 2005). For instance, Kapoor and Khan (2019) found that participants generated more feasible ideas to real-world tasks, while more original ideas were generated in object-based tasks, such as the Alternate Uses Task (AUT; Guilford, 1967). The main explanation is that realistic tasks elicit fewer original ideas, because the realistic nature of the problem pushes individuals to draw from their (easily) accessible experience and memory.

Another reason for the perceived negative correlation is that highly original ideas are more likely to be judged as less feasible because they involve, by definition, a step into the unknown (Baer, 2012). The most original ideas are often those that are radically different from existing solutions or practices, which often causes people to perceive them as less feasible. This will especially happen in real-life situations where people have to make a decision as to which idea to choose or support. In his influential paper about idea implementation, Baer (2012) wrote "ideas that are useful yet novel are likely to produce uncertainty and, as a result, are likely to be met with skepticism and hesitation […]. Thus […] the very nature of these ideas is likely to generate reluctance about their implementation" (p. 1103).

1.4 Theory: Four P's Framework

Rhodes (1961) suggested that creativity can be viewed as a property of a *person* (a trait or characteristic), *process* (a set of cognitive processes), *product* (a result of a process), or the environment (*press*). In this 4P's framework, *person* refers to the attitudes, dispositions, feelings, and beliefs within an individual (or group), *process* refers to the cognitive thinking processes that occurs when a person is engaged in a creative activity, *products* are the results of the creative activity, and *press* represents the setting or climate in which the creative person functions.

However, this theoretical framework has mostly been used in studying idea generation (or divergent thinking), while this thesis argues that a person's ability in recognizing and selecting creative ideas (i.e., *product*) depends strongly on both *how they think* (i.e., process), *who they are* (i.e., person), and this is subject to *environmental* or *situational constraints* (i.e., press). The available scarce research seems largely consistent with these proposition, and will be discussed in the following sections.

1.4.1 Product

As described earlier, definitions of what makes a product creative usually focus on originality and usefulness, also similarly termed appropriate, feasible, or effectiveness, among others (Horn & Salvendy, 2006, 2009; Zeng, Salvendy, & Zhang, 2009). How do we know whether people's judgment of their own and others' ideas for products can be considered accurate or inaccurate? Previous research commonly measured the accuracy of evaluations in terms of correctly identified creative ideas, so called 'hit rates' (see Runco & Smith, 1992; Wagner, 1993 for more information) or in terms of the discrepancy between participants and judges' evaluations (e.g., Amabile, 1982; Grohman et al., 2006) or assessed the extent of agreement in terms of the covariation between participant and judges' evaluations (see Silvia, 2008 for more information), or by means of informedness (see Benedek et al., 2016 for more information).

According to Amabile (1982), the assessment of creativity cannot be achieved by objective analysis alone (e.g., hit rates), but some types of subjective analysis is required. She argues that the concept of what is creative is largely shared among a domain's experts as tacit knowledge and that creativity should therefore be assessed by consensus between domain experts, also known as the consensual assessment technique (CAT; Amabile, 1982; 1996)[2]. If sufficient agreement was reached, this would define the level of the product's creativity, relative to the other products within a sample, within a particular context of time and place. In Grohman, Wodniecka, and Klusak's study (2006), after completing three divergent thinking tasks, people rated the creativity of each response on a 1 – 7 scale and estimated the percentage of other people who gave the responses. Several judges then rated the creativity of each response on a 1 – 7 scale, and the percentage of people who gave each response were computed. Accuracy was estimated by (1) the differences between self-ratings and the judges' ratings and (2) the differences between estimated percentages and observed percentages. For each idea, differences scores near zero represent agreement between self-ratings and judges' scores. Such a measurement has been used as a reliable measure for the degree of creativity associated with that product (Bødker & Iversen, 2002; Cseh & Jeffries, 2019).

Several studies have found that individuals are not good at *recognizing* and *selecting* creative ideas (Benedek et al., 2016; Blair & Mumford, 2007; Faure, 2004; Guo, Ge & Pang, 2019; Putnam & Paulus, 2009; Rietzschel et al., 2006; 2010; Simonton, 2003; Van Damme et al., 2019; Zhu et al., 2019). Benedek et al., (2016), for instance, found that people tended to underestimate the creativity of ideas. In this study, 1119 secondary teachers were presented with 72 ideas that were collected from various creative idea generation tasks (or divergent thinking tasks). The teachers were asked to decide which of those ideas are common, inappropriate or actually creative. They found a negative response bias which indicates a tendency to underrate the creativity of ideas, by judging creative ideas as common or inappropriate. Benedek et al., (2016) concluded that "people recognized that these creative ideas were original, but sometimes felt that they were not useful and appropriate enough to qualify as being creative" (p. 82). Another recent study found that highly original people, compared to less original people, tend to give lower ratings to other people's ideas (Guo et al., 2019). This tendency to underestimate others' creative ideas was more salient when they are instructed to pay attention to the novelty of the ideas or when the

2 For Amabile and her colleagues, the reliability and validity of the consensual assessment technique were conditional on following certain guidelines (Amabile, 1982, 1996; Hennessey, 1994; Amabile & Mueller, 2011): judges should (1) have domain experience; (2) rate creativity independently and subjectively, that is, without new training, discussion, or specific guidance; (3) rate creativity relative to a specific sample and context; (4) each see the items they are to rate in different random orders (to avoid order effects); and (5) when assessing a task for the first time, rate factors other than overall creativity (e.g., originality, quality) and to use factor analysis to ensure discriminant validity of the creativity measure.

novelty dimension outweighs the usefulness dimension in products. Furthermore, this tendency was even more apparent when the products being evaluated are the combinations of remote ideas/concepts, which suggests that perceived semantic distance mediates the effect of rater's originality on ratings of originality.

Several studies investigating idea selection found that people tend to reject highly original or risky ideas, and are more likely to select ideas that are consistent with social norms, and easy to understand (Blair & Mumford, 2007; Rietzschel et al., 2006; Rietzschel et al., 2010; Van Damme et al., 2019; Zhu et al., 2019). For instance, when instructed to select the *best* idea out of a pool of generated ideas, Rietzschel et al. (2006) found that groups of students did not select the ideas that received high creativity ratings by independent judges. Instead, students selected ideas that were higher in feasibility than in originality. Moreover, they found that the idea selection process was hardly more effective than taking a random sample of ideas. Another recent study confirmed the notion that students failed to select ideas that were more creative, original, and effective than the average idea (Zhu et al., 2019). They even found that students tended to select ideas that were less feasible than the average idea.

Little is known whether creativity in products are perceived similarly or different by people. While research on creativity differences among the two cultures of Creativity – *Art* and *Science* – has generally reported that art students are more open, report higher self-assessed creativity, and generate more ideas in divergent thinking tasks than science students, little is known whether creativity in products are perceived similarly or different by Art and Science students (e.g., Furnham & Crump, 2013; Hartlet & Greggs, 1997; Kaufman, Pumaccahua, & Holt, 2013a; Zare, 2011). Another limitation of these studies is that they define Science extremely broadly, lumping together everything from physics to medicine. As such, detecting where differences in *product* may emerge (or disappear) at a more specific level may be valuable for translating findings into useful guidance for educational practitioners (e.g., choice between engineering or mathematics). As such, Chapter 2 investigates whether there are differences in the perception of creativity in products within and between: (a) Art and Science; (b) Specific Science domains (STEM), and (c); Engineering studies (e.g., mechanical or electrical engineering). In line with prior research, creativity differences on person and process variables are investigated as well.

1.4.2 Process

It is commonly argued that idea evaluation is an *intuitive thinking process* (e.g., Petervari, Osman, & Bhattacharya, 2016; Stierand & Dörfler, 2016). However, "intuition or gut feeling" can be misleading, and more *analytical approaches* or *cognitive techniques* may improve people's ability in recognizing creative ideas (Basadur, Runco, & Vega, 2000; Eling, Griffin, & Langerak, 2016; Magnusson, Netz, & Wästlund, 2014; Stierand & Dörfler, 2016; Vernon, Hocking, & Tyler, 2016). For instance, Licuanan, Dailey, and Mumford (2007) have investigated whether active analysis of the original features of ideas would improve students' ability to recognize highly original ideas. In this study, 181 students attended a two-hours self-paced creativity training in which half the students were asked to describe the strengths and weaknesses of ideas with respect to the 'originality', while the other half was asked to provide a strengths and weaknesses analysis for the 'overall performance' of ideas. They found that this analytical approach or cognitive technique of strengths and weaknesses analysis improved people's ability to recognize highly original ideas. However, on the other hand, Ritter, Gu, Crijns, and Biekens (2020) found no effect of cognitive techniques on the ranking of new business ideas. In this study, 133 students attended a 140-hours cognitive-based creativity training where they learned to apply four cognitive techniques (i.e., simplify, differentiate, visualize, and tag the problem). In sum, there are mixed findings whether cognitive techniques improve people's ability to recognize creative ideas. As such, we aim to expand to this research field by investigating the effect of a cognitive-based creativity training on idea evaluation (i.e., Chapter 3).

In addition to cognitive techniques, having relevant *task exposure* or *familiarity* may help in recognizing both the novelty of a product and its usefulness. Several studies have investigated whether idea evaluation by the people who have generated the ideas (i.e., *intrapersonal* evaluation) is more accurate than idea evaluation done by people who have not generated the ideas (i.e., *interpersonal* evaluation). These studies have generally come to the conclusion that people are more accurate in evaluating their *own* ideas than *other's* ideas (Grohman et al., 2006; Runco & Smith, 1992; Silvia, 2008). Intrapersonal evaluation provides people with more task exposure or familiarity, and, as such, more insight into the associative history of each idea, and which ideas were rejected in favor of those that were kept. However, real-life settings are dominated by interpersonal evaluation, in which people have to judge ideas that were created by others without any task exposure (Berg, 2016). As such, we argue that task exposure may be key to enhance interpersonal accuracy of creative forecasting. Hence, in Chapter 4, we investigate the effect of task exposure on interpersonal idea evaluation.

1.4.3 Person

Some people are more open toward novelty than others, and may therefore be more motivated to recognize and select creative ideas. As such, *personality traits* might play an important role in idea evaluation and idea selection.

The relationship between personality and idea evaluation has not been extensively addressed (Basadur et al., 2000; Fürst, Ghisletta, & Lubart, 2016; Rodriguez, Cheban, Shah, & Watts, 2020; Stam, de Vet, Barkema, & De Dreu, 2013). Basadur et al. (2000) provide indirect evidence that a trait like *openness to experience* may impact idea evaluation. They found that the ability to produce a high number of ideas predicted both the originality of these ideas and the ability to accurately evaluate these ideas. The underlying predictor was a preference for avoiding premature convergence, in other words, the ability to postpone closure – a trait that highly correlates with openness (see also Stam et al., 2013). Furthermore, Fürst et al. (2016) found that idea evaluation is positively predicted by high scores on a cluster of personality characteristics called "convergence", which comprises characteristics such as persistence, precision, and critical sense, and correlates positively with *conscientiousness*.

Moreover, it seems plausible that personality might also play a role in idea selection. Toh and Miller (2016a) found that risk aversion – closely related to *openness to experience* – is significantly related to creative idea selection. They found that people who are more risk prone selected more creative ideas than people who are more risk averse. This result was confirmed in a team-level study as well (Toh & Miller, 2016b), where they found that teams who not only have higher levels of *tolerance of ambiguity* (closely related to openness to experience), but also higher levels of *conscientiousness* and *agreeableness* are more prone to select novel ideas.

In sum, the personality traits *openness to experience* and *conscientiousness* have been (indirectly) linked to idea evaluation, while the *openness to experience*, *conscientiousness*, and *agreeableness* have been linked to idea selection.

1.4.4 Situational constraints (press)

Cognitive thinking processes (e.g., cognitive techniques and task exposure), and our attitudes, dispositions, feelings, and beliefs (e.g., personality) are subject to environmental or situational constraints. Environmental or situational constraints are contextual factors that facilitate or inhibit creativity, such as rewards, surveillance, competition and expected evaluation (Amabile, 1982; 1984; Amabile, Hennessey, & Grossman, 1986; Wang, Wang, Liu, & Dong, 2017; Yuan & Zhou, 2002; 2008; Zhou & Oldham, 2001).

Many experimental studies have demonstrated negative effects of situational constraints on creativity (Amabile et al., 1986; Greene & Lepper, 1974; Lepper, Greene, & Nisbett, 1973; Loveland & Olley, 1979). In several studies, for example, children were asked to draw creative pictures under rewarded or non-rewarded conditions. It was generally found that children who expected reward performed less creatively than children who did not expect this (e.g., Amabile et al., 1986; Greene & Lepper, 1974).

However, more recent literature has found that different forms of situational constraints may have different effects on idea generation (e.g., Shalley & Perry-Smith, 2001). For example, it has been demonstrated that *controlling* aspects of contextual factors bring individuals' behavior under the control of the constraint and include some form of pressure to coerce a person to act in a specific way (e.g., provide participants with criticism on their ideas). In contrast, *informational* aspects of contextual factors provide people with information about their task competency and include behaviorally relevant information causing individuals to feel that they are performing competently on a task (e.g., provide participants with feedback on their ideas). Shalley and Perry-Smith (2001) confirmed this notion by asking students to generate solutions to a management problem with the expectation for an informational or a controlling evaluation. They found that students who expected a controlling evaluation generated less ideas than students who expected an informational evaluation.

As such, several researchers have addressed the effects of situational constraints on idea generation. The consequence is that we know quite a lot about the type or form of situational constraints and their effects on idea generation, but – given the lack of research on idea selection – still relatively little about the effect of situational constraints on idea selection. To our knowledge, there has only been one study that investigated the effect of a situational constraint on idea selection. Yuan & Zhou (2008) investigated the effect of the situational constraint - expected evaluation - on idea selection. For this purpose, they asked students not only to generate ideas, but also to select creative ideas for a management problem. During idea selection, students were reminded that their final solutions should be creative (i.e., both novel and appropriate). They found that students who expected evaluation modified their ideas more to make them appropriate and feasible to implement. This did not occur among students who did not have such an expectation. Moreover, students who expected evaluation of their ideas selected more creative ideas than students who did not expect evaluation.

In sum, while situational constraints have negative effects on idea generation, it seems plausible that these constraints have positive effects on idea selection. More research is needed on other situational constraints to confirm this notion. In Chapter 5, we aim to contribute to this research field by investigating the effect of another extrinsic constraints – expected implementation – on idea selection.

1.5 Aim of the thesis

The aim of the thesis is to investigate whether the *evaluation* of ideas and *selection* of ideas can be improved in educational settings. As described earlier, students' ability to recognize and select creative ideas (i.e., product) depends strongly on cognitive thinking processes (i.e., process), attitudes, dispositions, feelings, and beliefs (i.e., person), and is subject to situational constraints (i.e., press). As such, the 4 P's are used as theoretical framework. Figure 1.1 provides a conceptual overview of the four empirical studies included in this thesis. The thesis addresses the following research questions:

- Chapter 2: How do Art and Science students perceive creativity in products?
- Chapter 3: To what extent do cognitive techniques improve idea generation and idea evaluation skills in a cognitive-based training?
- Chapter 4: To what extent does task exposure to the idea generation process improve idea evaluation skills?
- Chapter 5: To what extent does expected implementation of ideas affect idea selection of novel and feasible ideas?

Figure 1.1

Conceptual overview of the thesis

	Creativity		»»»		Innovation
	Problem definition	Idea generation	Idea evaluation	Idea selection	Idea implementation
Product	~	√	X: *Ch. 2*	X	√
Process	~	√	X: *Ch. 3 & 4*	X	√
Person	~	√	X	X: *Ch. 5*	√
Press	~	√	X	X: *Ch. 5*	√

Note: the creative process from problem definition towards idea implementation is adapted from Amabile & Pratt, 2016; the four P's are adapted from Rhodes (1961); √ stands for a substantially amount of research, ~ for moderate amount of research, and X stands for a lack of research.

1.6 Thesis outline and main results

The 4P's provide a framework to explore one of the most long-standing debates whether creativity is domain-general or domain-specific (see for example, Plucker & Beghetto, 2004; Plucker et al., 2004). A large body of research has investigated this question by comparing Art and Science students on the 4Ps (e.g., Furnham, Batey, Booth, Patel, & Lozinskaya, 2011; Kaufman, Pumaccahua, & Holt, 2013b; Lubart & Guignard, 2004). Based on this comparison, they concluded that creativity is partly general, and partly domain-specific. However, *Science* has been defined extremely broadly in previous research, lumping together everything from physics to medicine. To gain more insight in the domain-generality domain-specificity debate, it is essential to zoom in as well on different Science domains to detect where differences in person, process and product may emerge or disappear. **Chapter 2** contributes to this debate by investigating creativity differences, and the magnitude of those differences, among students. More specifically, we examined differences in creativity within and between: (a) General Thematic Areas (Art and Science); (b) Specific *Science* domains (STEM), and; (c) Engineering micro-domains, for a total of 2277 students in German higher education institutions. A series of MANCOVA and ANCOVA analyses showed many statistically significant, but uniformly small, differences at all levels, across a range of person, process and product variables. The pattern of results suggests that openness, creative self-efficacy and divergent thinking may be general pre-requisites for creativity. In contrast, the way that characteristics of creative products (e.g., originality) are perceived appears more complex. The main finding is that students seem to differ in their perception of characteristics of creative *products* (e.g., originality, feasibility).

Chapter 3 investigates whether cognitive techniques in a cognitive-based training can improve idea evaluation skills among students (i.e., *process*). A pre-post-test within-subject design was conducted among 51 undergraduate students from a large university in the Netherlands. All 51 students received the 10-hour training as part of their bachelors programme, but were assigned to receive the training in the first or second semester. As such, students participated in both experimental conditions (control and intervention), albeit at different times (within-subject design). The Alternative Uses Task (AUT) and specially designed idea evaluation tasks were used before and after the training. In the idea evaluation task, students were asked to evaluate ideas on their originality and feasibility. Their ratings were compared with experts' ratings. The General Linear Model (GLM) for repeated measures indicated that students generated more (i.e., fluency) and different kinds of ideas (i.e., flexibility) after training. However, the results were non-significant. In line with recent research, the results suggested that training does not impact idea evaluation skills among students. This suggests that cognitive techniques might not be the way to improve people's ability to recognize creative ideas.

Chapter 4 investigates the effect of task exposure on idea evaluation, as task exposure provides people with more insight into the associative history of each idea, and what kind of ideas were rejected in favor of those that were kept (i.e., *process*). For this purpose, 1864 students in German higher education institutions evaluated ideas on their creativity, originality and feasibility. Their ratings were compared to experts' ratings. The students were randomly assigned to one of the following conditions: *task exposure* (i.e., in this condition they had to generate and evaluate ideas for the same task) or *no task exposure* (i.e., in this condition they had to generate ideas for a different task than the idea evaluation task). The results showed that task exposure improves students' ability to accurately recognize creative and original ideas, and their ability to discriminate between highly feasible and unfeasible ideas. As such, these findings suggest that task exposure is beneficial for creative idea forecasting. Together, these results highlight the importance to carefully reconsider whether or not people should be exposed to the task before evaluating other's ideas.

In order to move from creativity to innovation, it is of vital importance to not only be able to evaluate ideas, but also to select the most creative ideas. It is commonly assumed that successful idea evaluation enables students to select the most creative ideas for actual implementation. However, even though students may be better at recognizing which ideas are creative, the commonly situational constraint of *expected idea implementation* might affect idea selection (i.e., press). Yuan and Zhou (2008) have shown for another, but related constraint that undergraduates modified their selected ideas more to make them appropriate and feasible to implement when they *expected evaluation*. Another example is Sharma's (1999) finding that many creative ideas are generated within companies, but few reach the implementation phase. These results suggest that an older population, such as undergraduates or graduates, is heavily affected by the *expectation of idea implementation* while little is known how this situational constraint affects children's decision making.

As such, **Chapter 5** investigates the effect of expected implementation of ideas on children' selection of novel and feasible ideas (i.e., final product innovation). The selection of novel ideas requires a certain level of openness, therefore, we also investigate whether children' personality moderates this relationship (i.e., *person*). To this end, 403 grade-6 children (age 10-13) were asked to select two innovative ideas to improve the use of a stuffed toy elephant with or without the expectation to actually implement these ideas in the classroom. The results showed that children who expected implementation were less likely to select original ideas, but more likely to select feasible ideas than children who had no expectation to implement ideas. Moreover, implementation focused more on feasibility as compared to originality when selecting innovative ideas. The personality trait conscientiousness was found to moderate this relationship. Children with a high conscientiousness were found to select more feasible ideas even though they were instructed to select innovative ideas and did not expect idea implementation. Together, the results highlight the importance for educators to carefully consider whether or not practical components should be part of assignments, and to tailor instruction in assignments to the individual needs of children.

Chapter 6 concludes and provides a summary of the main results, contributions of this thesis, and provides avenues for future research. **Chapter 7** reflects on practical implications of the findings presented in this thesis.

2

Differences in creativity across Art and STEM students: We are more alike than unalike

Abstract[3]

The aim of the present research is to investigate creativity differences, and the magnitude and nature of those differences, among university students. More specifically, we examined differ-ences in creativity within and between: (a) General Thematic Areas (Art and Science); (b) Specific Science domains (STEM), and; (c) Engineering micro-domains, for a total of 2277 students in German tertiary institutions. The results showed many statistically significant, but uniformly small, differences at all levels, across a range of Person, Process and Product variables. The pattern of results suggests that Openness, Creative Self-Efficacy and Divergent Thinking may be general pre-requisites for creativity. In contrast, the way that characteristics of creative products (e.g. originality) are perceived appears more complex. This research sheds additional light on long-standing debates regarding domain-generality/specificity and creativity.

3 van Broekhoven, K., Cropley, D., & Seegers, P. (2020). Differences in creativity across Art and STEM students: We are more alike than unalike. *Thinking Skills and Creativity, 38*, 100707.

This chapter is joint work with David Cropley and Philipp Seegers, and is published at Thinking Skills and Creativity. We thank Rolf van der Velden, Lex Borghans, Barbara Belfi and colleagues at the University of South Australia for helpful comments and suggestions.

2.1 Introduction

In the era of the *Future of Work*, there is now broad recognition that creativity is a vital 21st century skill (e.g. Cropley, 2019; Cropley, Cropley, & Sandwith, 2017; World Economic Forum, 2016). This is true for elementary (Vincent-Lancrin et al., 2019) and secondary education (Anwar & Aness, 2012), as much as it is true for tertiary education (e.g. Cropley, 2015; Cropley et al., 2017). However, while both the development and the application of creativity in young children appears domain-general, does creativity also change *qualitatively* with age (Russ & Fiorelli, 2010: p. 237)? Tubb, Cropley, Marrone, Patston, and Kaufman (2020) found some evidence supporting a positive relationship between domain-specific mathematical creativity and grade level in high school students, but is this the case more broadly? Does creativity become more domain specific as individuals move through the educational continuum, and if so, how?

This developmental trajectory of creativity is mirrored by the Four-C Model (Kaufman & Beghetto, 2009). Finding a new way to tie your shoelaces may not be novel or even very effective for other observers, but for a child, it may represent an example of the "intrapersonal insights or interpretations" (Kaufman & Beghetto, 2009: p.4) inherent in *mini-c* creativity. In elementary school, it may be sufficient to know that a child is, broadly speaking, *creative*: they are good at divergent thinking, they are willing to take risks and are open to new experiences, using these qualities to solve everyday problems creatively. However, as an individual moves from elementary to secondary and tertiary education, their experience of (and their need for) creativity seems likely to shift from *purely* domain-general to include a more domain-specific element. Kaufman and Beghetto (2009: p.3) compared a shift from mini-c to *little-c* creativity with Amabile's (1996) componential model, highlighting the addition of *domain-relevant skills* and task motivation to more domain-general *creativity* skills. This would seem to suggest that a creative artist differs from a creative engineer only in terms of their contrasting *domain-relevant skills*. Both are good divergent thinkers, both are open to new experiences and tolerant of ambiguity, and both are risk-takers, but the artist is adept at perceiving shapes and colors, while the engineer is proficient in calculus and trigonometry. Is this pattern supported by empirical evidence? As the individual moves through school and into tertiary education, developing from mini-c to little-c creativity, does a shift from domain-general to domain-specific creativity occur, and what is the nature of this shift?

Rhodes's (1961) 4Ps: Person, Process, Product and Press (Environment) provide a framework for exploring domain-general and domain-specific conceptualizations of creativity across the educational continuum. If *creativity* is domain-specific, does the creative artist differ from the creative engineer in terms of *who they are* (i.e., person), *how they think* (i.e., process), and even *how they perceive creativity* in artefacts (i.e., product)? Furthermore, if they exist, at what level do these differences emerge in a hierarchy of domains?

The Amusement Park Theoretical (APT: Baer & Kaufman, 2005) model of creativity illustrates this hierarchy of domains with three levels, from very broad to very specific. First are *General Thematic Areas* (GTA), similar to Snow's (1959) two *cultures* of Art and Science. The APT model then divides these areas into *Domains* (e.g. Psychology or Mathematics) and then *Micro-domains* (e.g. Educational Psychology, Mechanical Engineering). A key concept of the APT model is that the Person, Process and Product factors of the 4P's framework (Rhodes, 1961) may give rise to separate patterns of individual differences across GTA, domains, and micro-domains. While there is evidence that this may hold true for GTA, and even domains, do these individuals differences persist at the level of micro-domains?

Without a clear, evidence-based understanding of the nature of creativity – domain-general or domain-specific – across the four elements of creativity (the 4Ps), it is hard to formulate strategies for nurturing specific creative competencies through high school and into universities, at the very time that students appear to transition from a domain-general form (i.e., mini-c) into a more domain-specific form of creativity (i.e., little-c). The aim of the present research is to investigate where, in the hierarchy of the APT model, differences emerge, and the magnitude and nature of those differences, with a particular focus on tertiary Science, Technology, Engineering and Mathematics (STEM) education.

2.1.1 General Thematic Areas (GTA) differences

There is already a considerable body of evidence regarding similarities and differences in the *General Thematic Areas* of creativity (e.g. Feist, 1998; Furnham & Crump, 2013; Kaufman & Baer, 2005; Kaufman et al., 2013b; Kaufman et al., 2016). This evidence, perhaps not surprisingly, suggests that there are some significant differences in personal factors (i.e., the Person), how they think when they are engaged in creative problem solving (i.e., the Process), and how they perceive creativity in artefacts (i.e., the Product).

A number of studies have focused, in particular, on differences in *Person* and *Process* factors between Arts and Science (Cattell, Eber, & Tatsuoka, 1970; Furnham & Crump, 2013; Grosul & Feist, 2014; Hartlet & Greggs, 1997; Kaufman et al., 2013b; Sagone & Caroli, 2012; Zare, 2011). For instance, Furnham and Crump (2013) found that art students, compared to science students, were more open, but less conscientious. Kaufman et al. (2013b) found that art students were less agreeable, but reported higher self-assessed creativity, than students in science studies (e.g. chemistry, mathematics, psychology). Moreover, Zare (2011) and Hartlet and Greggs (1997) found that art students generated more ideas than science students in divergent thinking (DT) tests (i.e., *fluency*). Sagone and Caroli (2012) found that boys attending arts schools, compared to boys attending science schools, scored higher on *elaboration* (i.e., the richness of detail in the ideas). We therefore hypothesized that, *on a cluster of Person (i.e., personality and creative self-efficacy) and Process (i.e., divergent thinking) variables, there will be statistically significant differences between Art and Science students* (H1a).

In addition, the present study also explored *Product* differences across GTA. Artists, on the one hand, create artefacts including, for example, sculptures, paintings, and drawings. For artists, *novelty* (or originality) and *aesthetic* qualities (e.g. beauty) are important in determining the success of their products. By contrast, science – or STEM disciplines more specifically – revolve around the development of (technological) solutions in response to an identified need. In this area, the *feasibility* of a solution, and its *effectiveness*, are as important as originality (see Cropley & Kaufman, 2012; Cropley, Kaufman, & Cropley, 2011 for a more detailed discussion). We therefore hypothesized that *Art students will associate novelty (originality) with creativity more strongly than do Science students (H1b), and, that Science students will associate feasibility and effectiveness with creativity more strongly than do Art students* (H1c).

2.1.2 Domain differences in STEM

Studies of GTA, not surprisingly, define *Science* extremely broadly, lumping together everything from physics to medicine. Detecting where differences in Person, Process and Product may emerge (or disappear) requires a shift deeper into the hierarchy of the APT model. Prior research has focused on specific domains such as medicine or economics (Eisenman, 1969; Lievens, Coetsier, De Fruyt, & De Maeseneer, 2002; Lounsbury, Smith, Levy, Leong, & Gibson, 2009; Marrs, Barb, & Ruggiero, 2007; Pringle, DuBose, & Yankey, 2010). Pringle et al. (2010), for example, compared eight different *business* domains (e.g. accounting, marketing, economics) and found that marketing majors were the most extraverted compared to any other major, with no significant differences in creativity across the business majors.

While there has long been recognition that the GTA of *Science* involves abundant creative thought (e.g. Curtin, 1982; Simonton, 2004, 2009; Tauber, 1996; Wechsler, 1979), key domains within this thematic area – Science, Technology, Engineering and Mathematics (STEM) – have been neglected by researchers (Dexter & Kozbelt, 2013; Kozbelt, Dexter, Dolese, & Seidel, 2012). An exception is Kline and Lapham (1992) who compared art, social science, science and engineering students on personality. They found significant differences not only between GTA (i.e., art versus science/engineering students), but also *between* the "science" domains, with (natural) scientists scoring higher on conscientiousness than engineers, who scored higher than social scientists. This relative lack of research regarding differences between STEM domains is important, given the fact that these occupations are expected both to grow strongly (Hawksworth, Berriman, & Goel, 2018; International Labour Organization, 2017) and are least likely to be replaced by automation in the era of the *Fourth Industrial Revolution* (Industry 4.0) and the *Future of Work* (Bakhshi, Downing, Osborne, & Schneider, 2017; Frey & Osborne, 2017). Drawing on the limited research, we hypothesized that*, on the cluster of Person (i.e., personality and creative self-efficacy) and Process (i.e., divergent thinking) variables, there will be statistically significant differences between students in the four STEM domains* (H2a).

Following the study of differences in GTA, we also explored *Product* differences between STEM domains. Natural sciences focus on knowledge development, while engineering and technology are concerned with needs-driven problem solving. Effectiveness and feasibility, as explained previously, are central to creativity in needs-driven problem solving. In contrast, knowledge creation in natural sciences may not be bound by the same imperative, with novelty and elegance, like the artist, possibly of greater importance as components of creativity. We hypothesized, therefore, that *natural science students will associate feasibility and effectiveness with creativity less strongly than technology and engineering students* (H2b). In addition, creativity in mathematics manifests itself in theoretical contributions where practical feasibility may play a less important role. In contrast, creative products in natural science, engineering, and technology must not only be original, but also feasible. Therefore, we expect that *mathematics students will associate feasibility with creativity less strongly than natural science, technology, and engineering students* (H2c).

2.1.3 Micro-domain differences in STEM

Notwithstanding any observed differences between GTA and STEM domains, detecting where differences in Person, Process and Product may arise between *micro-domains* is vital in translating findings into useful guidance for secondary and tertiary educators. Therefore, we also investigated differences between the micro-domains of *Engineering* (i.e., industrial, electrical, mechanical and civil engineering). In practical terms, more students might be torn between studying a micro-domain of engineering than students debating between pursuing a career in engineering versus art.

Only a handful of studies has examined creativity differences by micro-domains. Agogué, Le Masson, Dalmasso, Houdé, and Cassotti (2015) investigated how industrial designers and engineers perform on divergent thinking (DT) tests (specifically, on *fluency).* They found that industrial designers generated more ideas than engineers. Another recent study investigated whether ratings of product creativity differed between industrial designers and engineers (Cropley & Kaufman, 2019). They found that industrial designers rated the aesthetic quality of products higher than engineers, while they found no differences for the rating of the functionality of products.

These observed differences, however, may result from the fact that industrial design is distinctly different from engineering – i.e., they are *not* closely related micro-domains? Industrial design may be said to originate from art, while engineering is grounded in the physical sciences. The comparison may therefore be at the level of domain or even GTA, rather than of micro-domains. Comparisons of the more closely, and clearly, related micro-domains of engineering may provide more valuable insights for the debate on the domain generality or domain specificity of creativity. As such, the current study will compare four micro-domains of engineering: industrial, electrical, mechanical and civil engineering.

As related micro-domains, all are concerned with the development of technological solutions in response to an identified need. However, different engineering micro-domains might differ in the type of problems they solve, or even in the way they solve the same problem (Bruch & Krieshok, 1981; Veurink & Sorby, 2013; Veurink & Hamlin, 2011). For instance, electrical engineers design, develop, and test electrical power systems (e.g. motors) and electronic devices (e.g. microprocessors). Mechanical engineers design, develop, and test power-producing systems such as internal combustion engines. Civil engineers design, develop, and test structures in the built environment (e.g. bridges). Industrial engineering, somewhat in contrast, is concerned with the design, development and testing of *integrated systems* of people, materials, equipment, and energy. A difference between these closely related micro-domains might be that electrical, mechanical and civil engineers foster advances through innovative new technologies, while industrial engineers *apply* those new technologies (Kimbler, 1995).

Nazzal (2015) has investigated how civil, chemical, manufacturing, electrical, mechanical, transport systems and environmental engineering differed across the stages of the creative problem-solving process (i.e., problem recognition, idea generation, idea evaluation, and idea selection). The study found no differences between engineering micro-domains on idea generation and idea evaluation. However, students in chemical, environmental, and transport systems engineering selected more creative ideas than students in civil, mechanical, and manufacturing engineering.

Drawing on the limited available research, we hypothesized that*, on the cluster of Person (i.e., personality and creative self-efficacy) and Process (i.e., divergent thinking) variables, there will be statistically significant differences between four micro-domains of engineering (i.e., industrial, electrical, mechanical and civil engineering)* (H3a).

The present study also examined *Product* creativity differences between the micro-domains in question. The focus on the implementation and use of developed technologies in industrial engineering suggests that this micro-domain is more strongly user-centric than electrical, mechanical and civil engineering (Brawner, Camacho, Lord, Long, & Ohland, 2012; Kimbler, 1995). Consequently, it may be that *novelty* (originality) and *aesthetic* qualities (e.g. beauty), are more significant for industrial engineers, whose task is not only to implement a solution that works (effectiveness/feasibility), but one that meets underlying expectations for improvement or advancement (i.e., novelty), and is also complete, pleasing and user-friendly (i.e., elegant). We hypothesized that *industrial engineering students will associate novelty (originality) more strongly with creativity than electrical, mechanical and civil engineering students* (H3b).

2.2 Methods

2.2.1 Participants

A total of 2277 both undergraduate and graduate students (1052 females; 1225 males) aged between 17 and 37 (mean = 23.02; SD = 3.30) participated in this study. The students were enrolled in universities or universities of applied sciences[4] across Germany. 130 participants (91 females; 39 males) were enrolled in an Art degree, while 2147 participants (961 females and 1,186 males) were enrolled in a STEM domain (i.e., science, technology, engineering or mathematics).

Of the students enrolled in STEM domains, 665 participants (420 females; 245 males) were enrolled in a *natural science* degree (e.g. biology, physics); 461 participants (172 females; 289 males) were enrolled in a *technology* degree (i.e., computer science); 876 participants (280 females; 596 males) were enrolled in an *engineering* degree (e.g.

4 Universities are theory and research-oriented while universities of applied sciences are more practical and profession-oriented.

civil engineering, mechanical engineering), and 145 participants (89 females and 56 males) were enrolled in a *mathematics* degree.

In the Engineering micro-domains, 233 participants (91 females; 142 males) were enrolled in *industrial engineering;* 168 participants (37 females; 131 males) were enrolled in *electrical engineering*; 311 participants (76 females; 235 males) were enrolled in *mechanical engineering*, and 164 participants (76 females; 88 males) were enrolled in *civil engineering*.

2.2.2 Measures

The current study was part of the 12th round of the survey-based research project "Fachkraft 2030" which took place in March 2018. The survey is conducted by Maastricht University in cooperation with Studitemps GmbH. The survey consisted of questions about students' study experiences, participation in student jobs and future career expectations. At the end of the questionnaire, students could participate in a variety of psychological tasks (e.g. personality and creativity). Three Ps from Rhodes (1961) were measured in this study: *Person* (i.e., personality and creative self-efficacy), *Process* (i.e., divergent thinking), and *Product* (i.e., ratings of solutions on their creativity).

Person. Personality was measured using the 50-item version of the International Personality Item Pool (IPIP: Goldberg et al., 2006). For each personality trait, participants received ten statements (presented in a random order). Sample statements include: "I enjoy hearing new ideas" and "I am always prepared". Participants rated how well each statement describes themselves on a Likert scale, ranging from 1 (very inaccurate) to 5 (very accurate). Scale reliabilities (Cronbach's alpha) were strong: openness (α = .82), conscientiousness (α = .81), extraversion (α = .86), agreeableness (α = .77), neuroticism (α = .86). Second, creative self-efficacy (CSE) was measured using the three-item scale of Beghetto (2006). The three items are: "I am good at coming up with new ideas", "I have a lot of good ideas", and "I have a good imagination". Participants rated to what extent they disagreed or agreed with each item on a Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). Scale reliability (Cronbach's alpha) was good (α = .72).

Process. Divergent thinking (DT) was assessed by asking participants to generate ideas for a function-first or form-first divergent thinking task.[5] In total, six different

5 *Function-first* problems state the desired outcome, e.g. "How to transport baked beans?" for which solutions are sought. *Form-first* problems, in contrast, state the solution, e.g. "A Tin Can" and seek different possible uses. The former is more representative of real-world problem-solving (see Cropley, Cropley & Sandwith, 2017).

tasks were randomly distributed over the questionnaire. Each participant was asked to generate ideas for one divergent thinking task (no time limit). For the function-first task, participants were asked to generate as many ideas as they could for one of the following problems: 'How to improve the use of public trains', 'How to protect the environment', 'How to make waiting time at cash desks more bearable', or 'How can you make teaching in your university better' (see Baas, De Dreu, & Nijstad, 2011; de Buisonjé, Ritter, de Bruin, ter Horst, & Meeldijk, 2017; Ritter, van Baaren, & Dijksterhuis, 2012). For the form-first task, participants were asked to generate as many different and unusual uses as they could for a newspaper or a brick, as examples of the Alternate Uses Task (AUT: Guilford, 1967). Participants' responses were counted to produce a *fluency* score. Fluency has been used as a measure of creative potential in many studies (Batey, Chamorro-Premuzic, & Furnham, 2009; Batey & Furnham, 2008; Tsakanikos & Claridge, 2005). In addition, the average number of characters used per idea was counted to produce an *elaboration* score. Neither the flexibility nor the originality of ideas was measured: the total number of ideas generated (6,654) exceeding what could be rated in this study.

Product. Participants were presented with four solutions[6] for one of the following problems: (a) how to improve the use of public trains, or; (b) how to protect the environment (see Baas et al., 2011; de Buisonjé et al., 2017 for more information). They were first asked to rate the "overall creativity" of each solution on a Likert scale ranging from 1 (not at all creative) to 5 (very creative). Next, in a random order, participants were asked to rate the solutions according to the "originality", "feasibility", and "effectiveness" of each idea, also on Likert scales ranging from 1 (not at all original/feasible/effective) to 5 (very original/feasible/effective).[7]

6 Prior to this study, Dutch undergraduate students were asked to generate ideas to *improve the use of public trains* or to *protect the environment*. A list of 39 unique solutions for each problem were rated by creativity experts and resulted in a high interrater reliability between the experts (see Tables A2.1 - A2.2 in the Appendix).

7 Originality is the extent to which the solution is novel, out of ordinary. Feasibility is the ease with which the solution can be implemented (given the current context such as financial resources, infrastructure, time required, legal issues). Effectiveness is the extent to which the solution helps to solve the problem.

2

2.3 Results

Preliminary assumption testing was conducted to check for normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multicollinearity, with no serious violations noted. In addition, the inter-item correlations of the Product creativity differences were less than 0.8, suggesting that originality, feasibility and effectiveness measure distinct, but related constructs. Tables 2.1, 2.2 and 2.3 present descriptive data for the Person, Process and Product creativity measures respectively, by *GTA*, *Domain* and *Micro-domain*.

To test hypothesized differences between the GTA, Domains, and Micro-domains, on a cluster of creativity measures (i.e., Person and Process), this study used a series of one-way, between groups multivariate analyses of (co)variance (MANCOVA). Two groups of dependent variables were used: (a) *Person,* including openness, conscientiousness, extraversion, agreeableness, neuroticism, and creative self-efficacy, and; (b) *Process*, including divergent thinking (i.e., fluency and elaboration). To test hypothesis 1a, we conducted a MANCOVA analysis where the independent variable (factor) was GTA (art versus science). To test hypothesis 2a, we performed a MANCOVA analysis where the four domain categories were included as the fixed factor (i.e., STEM: natural science, technology, engineering and mathematics). For hypothesis 3a, we conducted a MANCOVA analysis where the four micro-domain categories were included as the fixed factor (i.e., industrial, electrical, mechanical and civil engineering). Following each MANCOVA analysis, a series of univariate, one-way ANCOVAs were performed in order to explore the results for the dependent variables separately, using a Bonferroni correction to control for potential Type I error (Table 2.1 and 2.2).[8] Age and gender were held constant in both the MANCOVA and ANCOVA analyses.

Next, a distinction was made for the function-first and form-first divergent thinking tasks to explore whether differences in divergent thinking (i.e., Process) emerge for both tasks, or only one of them (Table 2.2). To this end, MANCOVA and ANCOVAs were performed in which age and gender where held constant.

8 It can be argued that there are eight tests of significance (not counting those for the intercept and control variables), so p-values more than 0.05/8 = 0.0063 should be declared insignificant at the 5% level.

Table 2.1
Mean personality traits (Big 5) and creative self-efficacy (CSE) scores per GTA, Domain and Micro-domain

	N	Openness *M* (SD)	Conscientiousness *M* (SD)	Extraversion *M* (SD)	Agreeableness *M* (SD)	Neuroticism *M* (SD)	CSE *M* (SD)
GTA							
Art	130	3.82* (0.05)	3.33 (0.03)	3.11 (0.03)	4.15* (0.13)	2.91 (0.17)	4.03* (0.04)
Science	2147	3.54* (0.05)	3.42 (0.03)	3.05 (0.04)	3.86* (0.14)	3.12 (0.18)	3.70* (0.05)
Total	2277	3.56 (0.08)	3.42 (0.04)	3.05 (0.04)	3.88 (0.15)	3.11 (0.19)	3.72 (0.09)
Domain							
Science	665	3.61* (0.06)	3.36* (0.04)	3.07* (0.04)	3.97* (0.12)	3.04 (0.18)	3.67 (0.04)
Technology	461	3.51* (0.06)	3.36* (0.04)	2.89* (0.04)	3.77* (0.12)	3.11 (0.18)	3.70 (0.05)
Engineering	876	3.50* (0.06)	3.49* (0.04)	3.12* (0.04)	3.81 (0.11)	3.19 (0.17)	3.73 (0.04)
Mathematics	145	3.60 (0.06)	3.49 (0.04)	3.00 (0.04)	3.98 (0.12)	3.08 (0.18)	3.63 (0.05)
Total	2147	3.54 (0.05)	3.42 (0.03)	3.05 (0.04)	3.86 (0.14)	3.12 (0.18)	3.70 (0.05)
Micro-domain							
Industrial engineering	233	3.51 (0.06)	3.53 (0.02)	3.23* (0.06)	3.91* (0.08)	3.17 (0.19)	3.66 (0.05)
Electrical engineering	168	3.46 (0.05)	3.42 (0.02)	3.11 (0.06)	3.70* (0.08)	3.24 (0.16)	3.64 (0.04)
Mechanical engineering	311	3.52 (0.05)	3.49 (0.02)	3.05* (0.06)	3.76 (0.08)	3.19 (0.16)	3.81 (0.04)
Civil engineering	164	3.49 (0.05)	3.51 (0.02)	3.14 (0.06)	3.86 (0.09)	3.15 (0.19)	3.76 (0.04)
Total	876	3.55 (0.05)	3.42 (0.03)	3.06 (0.04)	3.82 (0.13)	3.16 (0.17)	3.71 (0.04)

Note: * Significant difference ($p < .0063$) between bolded asterisked score and asterisked score. CSE = Creative Self-Efficacy, M = Mean, SD = Standard Deviation. Age and gender were held constant.

Table 2.2
Mean divergent thinking per GTA, Domain and Micro-domain: General, Function- and Form-first divergent thinking tasks

		General			Function-first			Form-first	
	N	Fluency *M* (SD)	Elaboration *M* (SD)	N	Fluency *M* (SD)	Elaboration *M* (SD)	N	Fluency *M* (SD)	Elaboration *M* (SD)
GTA									
Art	130	4.13* (0.08)	49.41 (2.38)	88	3.74* (0.02)	59.12 (2.57)	42	4.91 (0.18)	28.34 (2.67)
Science	2147	2.62* (0.09)	41.34 (2.56)	1406	2.17* (0.02)	49.73 (2.78)	741	3.48 (0.20)	25.54 (2.87)
Total	2277	2.71 (0.36)	41.86 (3.17)	1494	2.26 (0.37)	50.27 (3.52)	783	3.56 (0.39)	25.70 (2.93)
Domain									
Science	665	3.08* (0.05)	44.64* (3.38)	444	2.58* (0.03)	54.19* (4.08)	221	4.10* (0.10)	25.33 (2.78)
Technology	461	2.34* (0.06)	44.1 (3.45)	293	1.79* (0.03)	52.75 (4.08)	168	3.30 (0.11)	28.81 (3.22)
Engineering	876	2.46* (0.05)	36.98* (3.41)	571	2.09* (0.03)	44.08* (3.95)	305	3.16* (0.10)	23.62 (2.82)
Mathematics	145	2.35* (0.06)	44.68 (3.41)	98	1.91 (0.03)	53.18 (4.15)	47	3.28 (0.11)	27.38 (3.10)
Total	2147	2.62 (0.09)	41.40 (2.58)	1406		49.73 (2.78)	741	3.48 (0.20)	25.54 (2.87)
Micro-domain									
Industrial engineering	233	2.39 (0.13)	41.41 (5.03)	165	1.93 (0.02)	47.81 (5.86)	68	3.44 (0.40)	26.09 (3.87)
Electrical engineering	168	2.27 (0.12)	34.46 (4.28)	107	2.09 (0.01)	40.20 (4.96)	61	2.57 (0.35)	24.30 (3.43)
Mechanical engineering	311	2.59 (0.12)	36.84 (4.29)	198	2.15 (0.01)	44.17 (5.00)	113	3.34 (0.34)	24.05 (3.38)
Civil engineering	164	2.52 (0.12)	33.55 (4.64)	101	2.21 (0.01)	42.06 (5.49)	63	3.10 (0.34)	19.56 (3.35)
Total	876	2.60 (0.08)	41.99 (2.43)	571	2.16 (0.02)	50.44 (2.61)	305	3.45 (0.19)	25.71 (2.77)

Note: * Significant difference ($p < .0063$) between bolded asterisked score and asterisked score. M = Mean, SD = Standard Deviation. Age and gender were held constant.

2

In addition to creativity differences in Person and Process, the present study also explored *Product* creativity differences across General Thematic Areas, Domains and Micro-domains. To test these hypotheses (H1b, H1c, H2b, H2c and H3b), a Fisher r-to-z-transformation was performed, assessing the significance of the difference between the correlation coefficients for the independent variables of GTA, Domain, and Micro-domain (Table 2.3 and 2.4).

Table 2.3

Correlations coefficients for creativity, originality, feasibility, and effectiveness per GTA, Domain and Micro-domain

		N		Originality		Feasibility		Effectiveness	
GTA	Art	130	Creativity	0.675	***	0.124		0.311	***
	Science	2147	Creativity	0.543	***	0.164	***	0.323	**
	Total	2277							
Domain	Science	665	Creativity	0.526	***	0.161	***	0.245	***
	Technology	461	Creativity	0.567	***	0.184	***	0.409	***
	Engineering	876	Creativity	0.547	***	0.180	***	0.340	***
	Mathematics	145	Creativity	0.512	***	-0.013		0.272	***
	Total	2147							
Micro-domain	Industrial engineering	233	Creativity	0.481	***	0.093		0.322	***
	Electrical engineering	168	Creativity	0.658	***	0.332	***	0.468	***
	Mechanical engineering	311	Creativity	0.567	***	0.179	**	0.341	***
	Civil engineering	164	Creativity	0.477	**	0.130		0.228	***
	Total	876							

Note: Significance levels indicated as follows: * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 2.4
Fisher r-to-z transformations per GTA, Domain and Micro-domain

		Creativity vs. Originality	Cohen's q	Creativity vs. Feasibility	Cohen's q	Creativity vs. Effectiveness	Cohen's q
GTA	Arts vs. Science	2.32 *	0.211	-0.45		0.93	
Domain	Science vs. Technology	-0.96		-0.39		-3.03 **	0.184
	Science vs. Engineering	-0.57		-0.38		-2.02 *	0.104
	Mathematics vs. Science	-0.22		-1.90		0.30	
	Mathematics vs. Technology	-0.21 *	0.078	-2.07 *	0.173	0.31	
	Mathematics vs. Engineering	-0.54		2.17 *	0.169	-0.83	
	Engineering vs. Technology	-0.50		-0.07		-1.39	
Micro-domain	Industrial vs. Electrical	-2.60 **	0.265	-2.47 *	0.252	1.70	
	Industrial vs. Mechanical	-1.36		-1.01		-0.24	
	Industrial vs. Civil	0.05		-0.36		0.99	
	Electrical vs. Mechanical	1.52		1.70		1.58	
	Electrical vs. Civil	2.44 *	0.146	1.93		2.49 *	0.152
	Mechanical vs. Civil	1.28		0.52		1.27	

Note: N art = 130, N science = 2147, N natural science = 665, N technology = 461, N engineering = 876, N mathematics = 145, N industrial engineering = 233, N electrical engineering = 168, N mechanical engineering = 311, N civil engineering = 164. Significance levels indicated as follows: * $p < .05$, ** $p < .01$, *** $p < .001$.

2

2.3.1 General Thematic Areas (GTA) differences

The one-way MANCOVA suggested that there was a significant difference between art and science students on the combined dependent variables, $F(1, 2{,}275) = 12.15$, $p < .001$, Wilk's Lambda = .95, partial $\eta^2 = .051$ (after controlling for age and gender). Hence, hypothesis H1a was supported.

To explore the results for the dependent variables separately, univariate one-way ANCOVAs were conducted. There were significant differences at $p < .0063$ for openness ($F(1, 2{,}255) = 36.47$, $p = .000$, partial $\eta^2 = .016$), agreeableness ($F(1, 2{,}255) = 14.71$, $p = .000$, partial $\eta^2 = .007$), creative self-efficacy ($F(1, 2{,}255) = 28.15$, $p = .000$, partial $\eta^2 = .012$), and fluency ($F(1, 2{,}255) = 32.24$, $p = .000$, partial $\eta^2 = .014$). An inspection of the mean scores (see Table 2.1 and 2.2) shows that art students reported higher levels, compared to science students, of: (a) openness (art: $M = 3.82$, $SD = 0.05$; science: $M = 3.54$, $SD = 0.05$); (b) agreeableness (art: $M = 4.15$, $SD = 0.13$; science: $M = 3.86$, $SD = 0.14$); (c) creative self-efficacy (art: $M = 4.03$, $SD = 0.04$; science: $M = 3.70$, $SD = 0.05$), and; (d) fluency (art: $M = 4.13$, $SD = 0.08$; science: $M = 2.62$, $SD = 0.09$).

To further explore hypothesis 1a, we also made a distinction for the function-first and form-first divergent thinking tasks. The one-way MANCOVA suggested that there was a significant difference between art and science students on the combined dependent variables (i.e., fluency and elaboration) for the function-first task: $F(1, 1{,}492) = 10.86$, $p < .001$, Wilk's Lambda = .97, partial $\eta^2 = .028$; and the form-first task: $F(1, 781) = 8.93$, $p < .001$, Wilk's Lambda = .96, partial $\eta^2 = .044$ (after controlling for age and gender). Consequently, univariate, one-way ANCOVAs showed that there was only a significant difference at the $p < .0063$ in fluency for the function-first task (see Table 2.2): Art students reported a higher level of fluency ($M = 3.74$, $SD = 0.02$) than science students ($M = 2.17$, $SD = 0.02$). However, it is important to note that the difference between art and science students on the form-first task was close toward significance ($p = .0068$).

To test the hypotheses concerning product differences and GTA (H1b and H1c), correlations (Table 2.3) suggested that art students associate originality with creativity ($r = .675$, $p < .001$) more strongly than do science students ($r = 0.543$, $p < .001$). This difference (Table 2.4) between art and science students is significant ($z = 2.32$; $p < .05$) with a small effect size (Cohen's $q = .21$). Hence, hypothesis H1b was supported. In comparison, science students associate feasibility ($r = 0.164$, $p < .001$) and effectiveness ($r = 0.323$, $p < .01$) with creativity more strongly than do art students ($r = 0.124$, $p > .05$, and $r = 0.311$, $p < .001$ respectively). However, these differences between art and science students were not significant (Table 2.4). Hence, hypothesis H1c was not supported.

2.3.2 Domain differences in STEM

The one-way MANCOVA with domains as the factor suggested that there were significant differences between the domains of natural science, technology, engineering and mathematics (i.e., STEM) on the combined dependent variables, $F(3, 2{,}143) = 12.61$, $p < .001$, Wilk's Lambda = .84, partial $\eta^2 = .057$ (after controlling for age and gender). Hence, there was support for hypothesis H2a.

To explore the results for the dependent variables separately, univariate one-way ANCOVAs were conducted. There were significant differences in openness at $p < .0063$ for science versus technology: $F(1, 2{,}123) = 11.45$, $p = .001$, partial $\eta^2 = .005$; and science versus engineering: $F(1, 2{,}123) = 20.40$, $p = .000$, partial $\eta^2 = .010$. In addition, there were significant differences in conscientiousness at $p < .0063$ for engineering versus science: $F(1, 2{,}123) = 22.41$, $p = .000$, partial $\eta^2 = .011$; and engineering versus technology: $F(1, 2{,}123) = 15.38$, $p = .000$, partial $\eta^2 = .007$. Moreover, there were significant differences in extraversion at $p < .0063$ for technology versus science: $F(1, 2{,}123) = 18.08$, $p = .000$, partial $\eta^2 = .008$; and technology versus engineering: $F(1, 2{,}123) = 28.45$, $p = .000$, partial $\eta^2 = 0.013$. Additionally, there were significant differences in agreeableness at $p < .0063$ for science versus technology: $F(1, 2{,}123) = 13.38$, $p = .000$, partial $\eta^2 = 0.006$. There were also significant differences in elaboration at $p < .0063$ for science versus engineering: $F(1, 2{,}123) = 12.94$, $p = .000$, partial $\eta^2 = 0.006$. Finally, there were significant differences in fluency at $p < .0063$ for science versus technology: $F(1, 2{,}123) = 18.19$, $p = .000$, partial $\eta^2 = 0.009$; and science versus engineering: $F(1, 2{,}123) = 19.49$, $p = .000$, partial $\eta^2 = .009$).

An inspection of the mean scores (Table 2.2 and 2.3) shows that natural science students reported higher levels, compared to technology students, of: (a) openness (science: $M = 3.61$, $SD = 0.06$, technology: $M = 3.51$, $SD = 0.06$); (b) extraversion (science: $M = 3.07$, $SD = 0.04$, technology: $M = 2.89$, $SD = 0.04$); (c) agreeableness (science: $M = 3.97$, $SD = 0.12$, technology: $M = 3.77$, $SD = 0.12$), and; (d) fluency (science: $M = 3.08$, $SD = 0.05$, technology: $M = 2.34$, $SD = 0.06$). In addition, natural science students reported higher levels, compared to engineering students, of: (a) openness (science: $M = 3.61$, $SD = 0.06$, engineering: $M = 3.50$, $SD = 0.06$); (b) fluency (science: $M = 3.08$, $SD = 0.05$, engineering: $M = 2.46$, $SD = 0.05$), and; (c) elaboration (science: $M = 44.64$, $SD = 3.38$, engineering: $M = 36.98$, $SD = 3.27$). Lastly, natural science students reported higher levels of fluency ($M = 3.08$, $SD = 0.05$) than mathematics students ($M = 2.35$, $SD = 0.06$). Engineering students reported higher levels, compared to technology students, of: (a) conscientiousness (engineering: $M = 3.49$, $SD = 0.04$, technology: $M = 3.36$, $SD = 0.04$), and; (b) extraversion (engineering: $M = 3.12$, $SD = 0.04$, technology: $M = 2.89$, $SD = 0.04$). In addition, engineering students reported higher levels of conscientiousness (M = 3.49, SD = 0.04) than natural science students (M = 3.36, SD = 0.04).

To further explore hypothesis 2a, we also made a distinction for the function-first and form-first divergent thinking tasks. The one-way MANCOVA suggested that there were significant differences between STEM domains on the combined dependent variables (i.e., fluency and elaboration) for the function-first task: $F(1, 1{,}402) = 14.81$, $p < .001$, Wilk's Lambda = .88, partial $\eta^2 = .046$; and the form-first task: $F(1, 737) = 6.91$, $p < .001$, Wilk's Lambda = .90, partial $\eta^2 = .041$ (after controlling for age and gender). Consequently, univariate one-way ANCOVAs showed that there were significant differences in fluency for both the function-first and form-first divergent thinking task (Table 2.2). Natural science students reported a higher level of fluency for function-first (M = 2.58, SD = 0.03) than technology (M = 1.78, SD = 0.03) and engineering students (M = 2.08, SD = 0.03). However, for the form-first task, natural science students (M = 4.08, SD = 0.09) reported a higher level of fluency only compared to engineering students (M = 3.16, SD = 0.10). In addition, natural science students (M = 54.20, SD = 4.11) only reported a higher level of elaboration for the function-first task than engineering students (M = 44, SD = 3.97).

To test the hypotheses concerning product differences and domains (H2b and H2c), correlations (Table 2.3) suggested that natural science students ($r = 0.161$, $p < .001$) associate feasibility with creativity less strongly than do technology ($r = 0.184$, $p < .001$) and engineering students ($r = 0.180$, $p < .001$). Neither the difference (Table 2.4) between natural science and technology students is significant, $z = -0.39$; $p > .05$, nor between natural science and engineering students ($z = -0.38$, $p > .05$). However, as expected, natural science students ($r = 0.245$, $p < .001$) associated effectiveness with creativity less strongly than technology ($r = 0.409$, $p < .001$) and engineering students ($r = 0.340$, $p < .001$). The difference between natural science and technology students is significant ($z = -3.03$, $p < .01$) as well as that between natural science and engineering students ($z = -2.02$; $p < .05$). However effect sizes for both (Cohen's q) were small ($q = .18$ and $q = .10$ respectively). Hence, hypothesis H2b was only supported for effectiveness.

In addition (Table 2.3) students in mathematics ($r = -0.013$, $p > .05$) associated feasibility with creativity less strongly than natural science ($r = 0.161$, $p < .001$), technology ($r = 0.184$, $p < .001$) and engineering students ($r = 0.180$, $p < .001$). The difference (Table 2.4) between mathematics and technology students is significant ($z = 2.07$; $p < .05$), as is that between mathematics and engineering students ($z = 2.17$, $p < .05$), with small effect sizes (Cohen's $q = .17$ for both). However, the difference between mathematics and natural science students was not significant (z = 1.90, p = .057). Hence, there was partial support for hypothesis H2c.

2.3.3 Micro-domain differences in STEM

The one-way MANCOVA with micro-domains as the factor suggested that there were significant differences between the micro-domains of Engineering (i.e., industrial, electrical, mechanical and civil engineering) on the combined dependent variables, $F(3, 872) = 2.69$, $p < .001$, Wilk's Lambda = .91, partial η^2 = .031 (after controlling for age and gender). Hence, there was support for hypothesis H3a.

To explore the results for the dependent variables separately, univariate one-way ANCOVAs were conducted. There were statistically significant differences in extraversion at the $p < .0063$ for industrial versus mechanical engineering: $F(1, 853) = 11.60$, $p = .001$, partial η^2 = .01; and in agreeableness for industrial versus electrical engineering: $F(1, 853) = 9.01$, $p = 0.003$, partial η^2 = .010. An inspection of the mean scores (Table 2.1) shows that industrial engineering students reported a higher level of extraversion (M = 3.23, SD = 0.06) than mechanical engineering students (M = 3.05, SD = 0.06). In addition, industrial engineering students reported a higher level of agreeableness (M = 3.91, SD = 0.08) than electrical engineering students (M = 3.70, SD = 0.08).

To further explore hypothesis 3a, we also made a distinction for the function-first and form-first divergent thinking tasks. The one-way MANCOVA suggested that there were significant differences on the combined dependent variables (i.e., fluency and elaboration) for the function-first task: $F(1, 567) = 2.61$, $p < .01$, Wilk's Lambda = .95, partial η^2 = .021; and the form-first task: $F(1, 301) = 2.25$, $p < .01$, Wilk's Lambda = .92, partial η^2 = .033 (after controlling for age and gender). However, these differences did not reach statistical significance according to the Bonferroni correction (p < .0063) in the univariate one-way ANCOVAs (see Table 2.2). Hence, hypothesis 3a was not supported when the distinction for function-first and form-first divergent thinking tasks was made.

To test the hypotheses concerning product differences and domains (H3b), correlations (Table 2.3) indicated that industrial engineering students (r = 0.481, p < .001) associate originality with creativity more strongly than do students in civil engineering (r = 0.477, p < .001), but less strongly than electrical engineering (r = 0.658, p < .001) and mechanical engineering students (r = 0.567, p < .001). Neither the difference (Table 2.4) between industrial engineering and civil engineering (z = -0.05, p > .05), nor that between industrial engineering and mechanical engineering (z = 1.36, p > .05) is significant. However, in contrast to our expectations, the difference between industrial engineering and electrical engineering is significant ($z = 2.60$, $p < .01$) with a small effect size (Cohen's $q = .26$). Hence, hypothesis H3b was not supported.

Table 2.5 presents an overview of the tested hypotheses in this study.

Table 2.5

Overview of the hypotheses in this study

		Hypotheses	p-value	Effect size	Supported/ Rejected	vs.
GTA	H1a	There will be *differences* between Art and Science students on a cluster of Person (i.e., personality and creative self-efficacy) and Process (i.e., divergent thinking) variables.	$p < .001$	$\eta^2 = .051$	Supported	
	H1b	Art students will associate novelty (originality) with creativity *more strongly* than do Science students.	$p < .05$	$q = .211$	Supported	
	H1c	Science students will associate feasibility and effectiveness with creativity *more strongly* than do Art students.	$p > .10$	-	Rejected	
Domain	H2a	There will be *differences* between students in the four STEM domains on a cluster of Person (i.e., personality and creative self-efficacy) and Process (i.e., divergent thinking) variables.	$p < .001$	$\eta^2 = .057$	Supported	
	H2b	Natural science students will associate feasibility with creativity *less strongly* than technology and engineering students.	$p > .05$	-	Rejected	*technology*
			$p > .05$	-	Rejected	*engineering*
	H2b	Natural science students will associate effectiveness with creativity *less strongly* than technology and engineering students.	$p < .01$	$q = .184$	Supported	*technology*
			$p < .05$	$q = .104$	Supported	*engineering*
	H2c	Mathematics students will associate feasibility with creativity *less strongly* than natural science, technology, and engineering students.	$p > .05$	-	Rejected	*natural science*
			$p < .05$	$q = .173$	Supported	*technology*
			$p < .05$	$q = .169$	Supported	*engineering*

2

Hypotheses			p-value	Effect size	Supported/ Rejected	vs.
Micro-domain	H3a	There will be *differences* between students in four micro-domains of engineering (i.e., industrial, electrical, mechanical and civil engineering) on a cluster of Person (i.e., personality and creative self-efficacy) and Process (i.e., divergent thinking) variables.	$p < .001$	$\eta^2 = .031$	Supported	
	H3b	Industrial engineering students will associate novelty (originality) *more strongly* with creativity than electrical, mechanical and civil engineering students.	$p < .01$	$q = .265$	Opposite direction	*electrical*
			$p > .05$	-	Rejected	*mechanical*
			$p > .05$	-	Rejected	*civil*

2.4 Discussion

2.4.1 Person

In line with previous research, the results of this study indicate clear differences at the level of General Thematic Areas (GTA). Art students are different from Science students in terms of personality. Broadly speaking, Art students are more open to new experiences, more agreeable (see also Feist, 1998) and have a higher Creative Self-Efficacy (CSE: see also Furnham et al, 2011) than their Science counterparts. Importantly, however, effects sizes in the present study suggest that these differences, while statistically significant, are *small*. Coupled with the present study's finding of *no* statistically significant differences on extraversion, conscientiousness and neuroticism at the level of GTA, this supports the notion of elements of Person as *general pre-requisites for creativity* in any area (Baer & Kaufman, 2005; Kaufman & Baer, 2005). Openness and CSE, therefore, support creativity for the Artist as much as they do for the Scientist. At this level (GTA) and for this factor (Person), educational support for creativity should foster openness and CSE as broadly as possible, from Kindergarten through to University education. It is important to note here that we confine our discussion to those aspects of the Person known to correlate to creativity, i.e., openness and creative-self efficacy (e.g. Batey & Furnham, 2006; Tierney & Farmer, 2002). Although the current study indicated a statistically significant difference in agreeableness at the level of GTA, the lack of any association of this factor with creativity (Batey et al., 2009) places it outside the scope of this discussion.

The current interest in STEM education, the *digitalization* at the heart of Industry 4.0, and the focus on key skills as an element of the Future of Work, provides a rationale for turning attention specifically to the *domains* that comprise STEM. The APT model suggests that if elements of personality are general pre-requisites for creativity across thematic areas, then they should remain general pre-requisites at the more specific level of domains. Openness, in other words, should remain a correlate for creativity for natural scientists, technologists, engineers and mathematicians, as much as it is a general pre-requisite for artists.

In the present study, some relevant personality differences are apparent among these four thematically related domains. Natural science students, for example, are more open to new experiences than technology and engineering students. Engineering students were more conscientious than natural science and technology students (we introduce this factor at the level of domains because there is support (Batey & Furnham, 2006) for its association with creativity). However, these significant differences were even smaller in magnitude than those observed at the level of GTA, supporting the role of openness and creative self-efficacy as *general* pre-requisites for creativity. This finding then adds weight to the importance of broad educational support for these Person factors of creativity, from Kindergarten to University.

Although the results for GTA and the STEM domains suggest that any difference on the factor Person diminishes the deeper we proceed into this hierarchy, the present study, nevertheless, compared four Engineering *micro*-domains (i.e., industrial, electrical, mechanical and civil engineering). No statistically significant differences on openness, conscientiousness, neuroticism or CSE were observed at this level. The small differences between artists and scientists became even smaller between STEM domains, and disappeared completely, in statistical terms, at the level of engineering micro-domains. Openness and CSE remain *general pre-requisites for creativity*. This result adds further weight to the need for broad educational support, now more specifically at tertiary level, for the Person factors that are linked to creativity.

2.4.2 Process

As with Person, hypothesized differences in Divergent Thinking (DT) – the core cognitive *process* associated with creativity – were found between the GTA under investigation. Art students, in this case, showed statistically significant, higher levels of fluency, both in general, and for function-first problems. However, as was the case for openness and CSE (i.e., Person factors), these differences remain small as determined by effect size. Coupled with the fact that there was no statistically significant difference between Art and Science students on elaboration, a case can be made that Divergent Thinking is also a *general pre-requisite for creativity*. The results for DT on the function-first and form-first tasks support this conclusion, with a statistically significant, but small, difference between art and science students only on fluency (Function-First task).

Turning to the more specific STEM domains, statistically significant differences emerged on general fluency between natural science students and technology, engineering, and mathematics students. For elaboration, a statistically significant difference was found between natural science and engineering. Differences were also noted for the function-first and form-first DT tasks. In all cases, however, the differences showed effect sizes even smaller than those between GTA. In practical terms, therefore, these differences were not meaningfully present between STEM domains, further supporting a divergent thinking as a *general pre-requisite for creativity*.

Proceeding to the level of micro-domains, the present study found no statistically significant differences on fluency or elaboration either in general, or for the function- and form-first tasks. Whether industrial engineer, electrical engineer, mathematician, natural scientist or artist, divergent thinking – a defining element of creativity – appears to be present to a statistically similar degree. The implications for elementary, secondary and tertiary education are therefore similar to those discussed for openness and CSE. Divergent thinking is a key *domain-general* creativity skill.

2.4.3 Product

In addition to differences on Person and Process, this study explored differences in how individuals perceive the qualities of a creative product across GTA, domains and micro-domains. Specifically, the degree to which individuals associate a product's originality, feasibility and effectiveness with creativity were examined.

At the level of GTA, both art and science students strongly associated originality (novelty) with creativity, with correlation coefficients >.500. This strong association was greater among art students compared to science students, with the difference in correlation coefficients statistically significant. The magnitude of this difference, however, was small. Both art and science students moderately associated effectiveness with creativity, with no statistical difference between these thematic areas. Feasibility was weakly associated with creativity, also with no statistical difference between art and science students. Broadly speaking, and consistent with the findings for Person and Process, there is little practical difference in the way art and science students perceive product creativity or associate key product qualities with creativity. Both, it can be said, see originality as central to defining the creativity of a product, with effectiveness also moderately important. This is consistent with mainstream views in creativity research (e.g. Sternberg, Kaufman, & Pretz, 2002; Cropley & Kaufman, 2012).

At the level of STEM domains, this pattern of associations remains the same. The four STEM domains associate originality with product creativity strongly, with some statistically significant, but small, differences between domains. Effectiveness has a moderate association with product creativity, also with some significant, but small, differences. Finally, feasibility is weakly associated with product creativity, with the exception of mathematics students. Mathematicians indicated no association between feasibility and product creativity, although the magnitude of the difference between them and the other STEM disciplines was very small. Notwithstanding the small effect sizes for the inter-domain differences, engineering and technology students in general indicated stronger associations of all three product qualities with creativity, compared to natural science and mathematics students.

Finally, differences were found between engineering micro-domains. Electrical engineering shows the strongest associations of originality, feasibility and effectiveness with creativity, followed by mechanical engineering, with civil engineering and industrial engineering somewhat weaker. However, all differences that are statistically significant are small in magnitude.

Two features of these results stand out in comparison to the results for Person and Process. First, unlike Person and Process, differences that were manifest at the level of GTA did not disappear the deeper the analysis proceeded into the hierarchy. Second, the differences detected both broadened as the analysis drove deeper into the hierarchy, and in some cases, enlarged. In plain language, a small difference in the association of only originality with creativity at the level of GTA grew into small associations of originality, feasibility and effectiveness at the level of domains, and indeed grew into slightly larger differences across all three factors at the level of micro-domains. Differences in openness and CSE (Person) and divergent thinking (Process) started small and stayed small, or vanished, whereas differences in the association of Product qualities with creativity started small but expanded slightly in both scope and magnitude.

2.4.4 Limitations and future research

Several limitations of this study should be noted. First, differences in both gender and age were not considered. As Baer and Kaufman (2008) make clear, there remain many questions about gender differences and creativity. The present study controlled for both age and gender, in order to focus on potential differences by GTA, domain, and micro-domain. Future research, however, should explore the potential impact of these factors, particularly in micro-domains.

Second, while person, process and product factors of the 4P's framework were investigated, we did not address the *press* (the *environment*) in which students operate. The fact that participants were drawn from many different tertiary institutions precluded any meaningful analysis of the impact of their learning environment in this study. It should be noted, however, that the finding – that individuals' assessments of product creativity are domain specific – likely stems from the contextual (i.e., press) elements associated with each domain (e.g. the different reward systems, demands and constraints present in each domain). The likely impact of press emerges, at the level of micro-domains, as differences in the norms and the *culture* of each discipline. Future studies therefore should investigate more explicitly possible differences between domains and micro-domains driven by specific environmental or contextual factors unique to those areas of activity. In simple terms, do engineers, for example, learn to *think like* engineers, in contrast to scientists or mathematicians? Does this then influence how these domains see creativity in products? Cropley and Kaufman (2019) explore this issue in relation to engineers and industrial designers.

A third limitation of this study relates to the variables chosen for analysis. Openness and creative self-efficacy have well-established links to creativity and are a logical starting point for an exploration of domain differences and creativity. With these relationships becoming clearer, an opportunity arises in future research to expand the range of variables examined. To what extent do other aspects of the person, for example persistence, grit or optimism, explain differences by domains? Do factors such as mood or emotional state help to explain the relationship between creativity, GTA, domains and micro-domains?

2.4.5 Conclusion

The results for Person and Process are indicative of factors that are general pre-requisites for creativity. Any individual, to be creative, benefits from high openness, high CSE, and a strong ability to think divergently. Education at all levels must respond accordingly, providing broad support for these elements of creativity. However, the results for Product suggest, albeit only weakly, that individuals' assessments of product creativity, and the relative importance of originality, feasibility and effectiveness to creativity, are domain specific.

These results support the notion (Plucker & Beghetto, 2004) that creativity, as a manifestation of who we are, and how we think, is general in nature. People who are open, flexible and adept at thinking divergently are best placed to be creative, and education systems at all levels should foster those qualities. Conversely, while all areas of endeavor recognize creativity in outcomes (products) as inseparable from originality and relevance/effectiveness, there are *discipline* specific differences in exactly how these qualities are valued. It is no surprise that engineers have a more *functional* (see Cropley & Cropley, 2005) view of product creativity – valuing effectiveness and feasibility in particular – whereas artists place greater emphasis on originality. Creativity in people is broadly domain general, but creativity in products is shaped by the needs, standards and cultures of the disciplines that produce those creative outcomes.

2.5 Appendix

Table A2.1

List of 39 ideas for 'how to improve the use of public trains?'

1	Audio-guide, which explains things about the surroundings.
2	Psychologist that is available in the train.
3	More suitable chairs and tables for working in and on.
4	Computer screens in the train seats.
5	Changing the train tracks into a rollercoaster.
6	Having trains drive until peoples' front doors.
7	Provide a more pleasant exterior and interior, designed by artists.
8	Improve the positive mood by distributing candy.
9	Train coupé for smokers.
10	Food and beverage dispenses in the train.
11	A bookshelf in every train.
12	Adding a points saving system to the public transport card.
13	A glass roof in the train.
14	Bar in the train.
15	A first aid post in the train.
16	A domestic cat in every train.
17	Increasing the number of trains and train stations.
18	Deploying extra night trains.
19	Breakfast service in the train.
20	Good toilets and bathrooms in the train.
21	'I-train app' to buy tickets, check in and receive personal travel info.
22	Placing alarm clocks in the train.
23	Childcare in the train and at the station.
24	Free food and drinks in the train.
25	Free travelling when all seats are taken.
26	Free WIFI in the train.
27	Increasing legroom and number of seats.
28	Organise an Open Day of Deutsche Bahn.
29	Involve the train passengers in designing the train.
30	Create faster trains.
31	Social media campaigns about train travel.
32	Speed-dating sessions in the train.
33	Improving transfer time and decreasing delay.
34	Heat rail changes better.
35	Less different train classes.
36	Decorate the train and dress up the conductors during the holidays.
37	Making train travel cheaper or for free.
38	Addition of train compartments for bike storage.
39	Create aesthetically pleasing stations.

Table A2.2

List of 39 ideas for 'how to protect the environment?'

1	Fighting deforestation.
2	Replacement of old buildings for energy-efficient new buildings.
3	Environmental protection as a school subject for everyone.
4	Producing and buying local food.
5	Use people to create natural energy themselves (e.g. cycling).
6	Use pictures of cute animals to emphasize the sadness of the meat.
7	Abolish bio-industry.
8	Treaty against environmental pollution.
9	Taxes on meat.
10	Increase the tax on meat.
11	More monitoring against poachers.
12	Organise a cooking evening to raise awareness of vegan cooking.
13	Taxes on wood.
14	Everyone must become vegan.
15	Give everyone a weapon.
16	Biodegradable products.
17	Better connections by public transport between town and village.
18	Tax fast food to subsidize the production of organic food.
19	Vegan Monday.
20	Letting people work much more at home.
21	Prohibiting growing avocados.
22	Develop new cleaner fuel for flying.
23	Design plastic that dissolves in salt water.
24	Create an ecological footprint per company and publish it.
25	Less smoking.
26	Top 500 list of greenest companies, for a marketing incentive.
27	More promotion of foodbank.
28	Sexual education in Africa.
29	Make more things from cork.
30	Make it possible to hand in reusable plastic to the supermarket.
31	More supermarkets without packaging.
32	Taxing wealthy people more severe.
33	New furniture made from waste products.
34	Reuse plastic bags for furniture.
35	Encourage space travel.
36	Remove oil and plastic from the sea.
37	Expand solar energy.
38	Television advertisements on the environmental problem.
39	Obligation to buy any solar system.

3

Fostering university students' idea generation and idea evaluation skills with a cognitive-based creativity training

Abstract[9]

This paper examines the effectiveness of a 10-hour cognitive-based creativity training on idea generation and idea evaluation among 51 undergraduate students (mean age 22) from a large university in the Netherlands. A pre-post-test within-subject design was conducted. The Alternative Uses Task (AUT) and specially designed idea evaluation tasks were used before and after the training. In the idea evaluation task, students were asked to evaluate ideas on their originality and feasibility. Their ratings were compared with content experts' ratings. General Linear Models (GLM) for repeated measures were conducted to determine whether any change in idea generation and idea evaluation is the result of the interaction between type of treatment (i.e., intervention or control group) and time (pre- and post-test). The results indicated that students did not generate significantly more (i.e., fluency) and different kind of ideas (i.e., flexibility) after training. Most importantly, in line with recent research, the results suggested that training does not impact idea evaluation

9 van Broekhoven, K., Belfi, B., Hocking, I., & van der Velden, R. (2020). Fostering University Students' Idea Generation and Idea Evaluation Skills With a Cognitive-Based Creativity Training. *Creativity. Theories–Research-Applications*, 7(2), 284-308, 102478.

skills among students. The implications of these results for educational practice and future research are discussed.

This chapter is joint work with Barbara Belfi, Rolf van der Velden, and Ian Hocking, and is published at Creativity. Theories – Research – Applications. We thank Joost Veth, Rose Vincent, Danilo Spinola, Joseph Park, Patima Chongcharoentanawat, Sanne van Wetten, Sander Dijksman, and the participants of the Creativity, Innovation and Entrepreneurship seminar 2018 in Maastricht, and the 2018 UK Creativity Researchers' Conference in Preston for their valuable comments and suggestions on earlier versions of this paper.

3.1 Introduction

One skill that is believed to be particularly important in the future labor market is creativity. This is because, up till now, computers are still not able to generate original and feasible ideas for complex problems, such as social, economic, and technological challenges (Autor, Levy, & Murnane, 2003). Therefore, policy makers and business leaders around the world have stressed that creativity should be fostered among graduates (Cachia, Ferrari, Ala-Mutka, & Punie, 2010; IBM, 2010). Research has further shown that creative thinking skills can be trained (see Scott, Leritz, & Mumford, 2004a for a meta-analysis). However, to date, creativity training is often not an integral part educational systems; in fact, the education system often discourages it (Baepler, Walker, & Driessen, 2014; Edwards & McGoldrick, 2006).

Creativity can be defined as "the interaction among aptitude, process, and environment by which an individual or group produces a perceptible product that is both novel (original, new) and useful (appropriate, feasible) as defined within a social context" (Plucker, Beghetto, & Dow, 2004, p. 90). *Idea generation* is a vital stage in this "interaction among aptitude, process, and environment" and depends heavily on divergent thinking capacity, that is, one's capability to generate numerous ideas (Kaufman & Sternberg, 2019). However, also a certain level of judgment is involved in creativity: it is not only important to have many ideas, but it is also vital to know which ideas are the most original and useful. As such, comprehensive theories of creativity suggest that in addition to idea generation, also *idea evaluation* is a vital process in the creativity process. This stage relies predominantly on convergent thinking capacity (i.e., one's capability to recognize the most original and feasible ideas, see Cropley, 2006; Fogler, LeBlanc, & Rizzo, 1995; Scott, Leritz, & Mumford, 2004b). The most common view on how divergent and convergent thinking styles are related to one another is that some moderate level of convergent thinking is necessary to be able to come up with many ideas (Kaufman & Sternberg, 2019). For example, Basadur, Runco, and Vega (2000) found that people's ability in idea generation was an important predictor for their ability to accurately recognize original ideas (i.e., idea evaluation).

The most popular way to enhance creativity has been training (Birdi, 2016; Scott et al. 2004a; Valgeirsdottir & Onarheim, 2017). As such, many organizations have invested substantial time and resources in the development and deployment of creativity training among their staff. For example, the Community Innovation Survey (CIS) – a questionnaire developed by Eurostat to investigate organizations' innovation activities in 19 countries – revealed that in 57% of firms that they deemed as ‘innovative’, engaged in innovation-related training activities (OECD, 2010). Also in higher education, creativity training - executed as either distinct course segments or workshops - has been developed for almost every student population including, for example, psychology students (Vally et al., 2019), nursing students (Liu, Wang, Chen, & Chao, 2020), engineering students (Zhou, 2012), and business students (Ritter, Gu, Crijns, & Biekens, 2020).

Since there is no established strategy yet regarding to how creativity can best be trained, creativity training programmes come in many shapes and forms. Scott et al. (2004a), carried out a meta-analysis study to categorize creativity training programmes as to whether or not they stressed (a) cognitive models, (b) social models, (c) personality models, (d) motivational models, (e) confluence models (supplemented cognitive models), or (f) other models (e.g. attitudes, blocks to creative thinking). They found that training programmes that involved cognitive skills were most effective (Scott et al. 2004a). A typical cognitive-based training programme focuses on various cognitive strategies for performing creative tasks, such as brainstorming or analogical thinking.

However, a limitation of many cognitive-based training programmes is that they predominantly focus on training in divergent thinking and ignore convergent thinking, because they assume that convergent thinking automatically occurs within the context of creative though, which has been found to be untrue (Scott, Leritz & Mumford, 2004b; Fasko, 2001). Therefore, Mumford, Baughman, and Sager (2003), among others, have argued that it is important to integrate both divergent and convergent thinking as principal components in cognitive-based creativity training programmes.

To date, there are only a few studies that have investigated the effect of a cognitive–based creativity training on divergent and convergent thinking (Basadur et al., 2000; Ritter & Mostert, 2016; Ritter, Gu, Crijns, & Biekens, 2020; Runco & Basadur, 1993). The results of these studies are mixed, where some studies reported positive results, other studies did not show an improvement in convergent thinking after training (see ‘past studies of cognitive-based training programmes’ for more information). For example, Basadur et al. (2000) found that managers recognized more accurately the originality of their ideas after a cognitive-based creativity training, wherein they experienced three stages of creativity (i.e., problem finding, idea generation and idea evaluation). In contrast, Ritter et al. (2020) found no effect of a similar cognitive-based creativity training on undergraduates’ ability to recognize more accurately which ideas were creative. To shed more light on this debate, the present study designed and tested a cognitive-based creativity training for undergraduates to enhance their skills in idea generation (i.e., divergent thinking) and idea evaluation (i.e., convergent thinking).

3.2 Past studies on cognitive-based training programmes

Numerous studies have investigated whether cognitive-based training may be a viable way to enhance divergent thinking skills (e.g. Abraham et al., 2019; Castillo, 1998; Hudgins & Edelman, 1988; Jausovec, 1994; Khatena, 1971; Ritter & Mostert, 2016; Ritter et al., 2020; Sun et al., 2019). For instance, in the study of Khatena (1971), 188 preschool children received a 6-h training which incorporated three creativity techniques (i.e., breaking away from the obvious and commonplace, restructuring, and synthesis). They found that training improved performance on divergent thinking (DT) tests (specifically, on *fluency*, *flexibility*, *originality*, and *elaboration*). Similarly, in the study of Sun et al. (2019), fifty undergraduates received a demonstration of a computer-based cognitive mapping tool for applying creativity techniques, whereas fifty other undergraduates did not receive this demonstration. They found that the computer-based cognitive mapping tool improved students' performance on divergent thinking (DT) tests (specifically, on *fluency*, *flexibility* and *originality*).

However, the majority of cognitive-based training programmes has only measured divergent thinking. Although important, divergent thinking is only one component of creative thinking. There have been fewer studies investigating the effect of cognitive-based training programmes on divergent as well as convergent thinking (Basadur et al. 2000; Birdi, 2007; Ritter & Mostert, 2016; Ritter et al. 2020; Runco & Basadur, 1993). The results of studies on the effect of creativity training on convergent thinking are mixed. While some studies reported positive results, other studies did not show an improvement in convergent thinking after training.

For example, in the study of Runco and Basadur (1993), thirty-five managers attended a 20-h cognitive-based creativity training wherein they experienced three stages of creativity on managerial problems (i.e., problem finding, idea generation and idea evaluation). In the training, the participants learned to apply various creativity techniques (e.g. brainstorming). Before and after the training, idea generation and idea evaluation tasks were conducted to measure divergent and convergent thinking (see Runco & Basadur, 1993 for more information). In the idea evaluation task, managers rated their own ideas on originality and their ratings were compared with the statistical infrequency of ideas to determine accuracy. The results suggested that training improved managers' ability to generate more ideas (i.e., *fluency*) and recognize more accurately the originality of their ideas. A subsequent study from Basadur et al. (2000) reported similar findings among 112 managers who recognized more accurately the originality of their ideas after training.

In contrast, Ritter et al. (2020) found no effect of creativity training on undergraduates' ability to recognize more accurately which business ideas were creative (i.e., convergent thinking). In this study, hundred thirty-three undergraduates attended a 140-h cognitive-based creativity training wherein they experienced six stages of creativity on a wide range of problems (i.e., understanding the question, convergent thinking, divergent thinking, detached thinking, stop thinking and sleeping). In the training, the participants learned to apply four creativity techniquess (i.e., simplify, differentiate, visualize, and tag the problem). Before and after the training, divergent and convergent thinking was measured. Divergent thinking was measured by a visual imagination task and the Alternative Uses Task (AUT: Guilford, 1967). Convergent thinking was measured by a convergent visual task, the Remote Associate Test (RAT: Mednick, 1962), and an idea selection task (see Ritter et al. 2020 for more information). The idea selection task is comparable to the idea evaluation task of Runco and Basadur (1993). In this task, participants had to rank order three pictures of business ideas from *most creative* to *least creative*. These business ideas had been evaluated by creativity experts as well to determine their accuracy. In line with Basadur et al. (2000) and Runco and Basadur (1993), they found that creativity training improved students' ability to generate more ideas (i.e., fluency) and their cognitive flexibility. However, in contrast, they found no effect of creativity training on any of the convergent thinking measures.

In sum, previous research has found that cognitive-based creativity training can enhance divergent thinking, but there are mixed findings concerning convergent thinking. As such, the first aim of the present study is to replicate existing findings regarding the effect of creativity training on idea generation (i.e., divergent thinking). The second aim is to contribute to the debate whether convergent thinking is a skill that can be enhanced via training. For this, we will investigate the effect of creativity training on idea evaluation (i.e., convergent thinking).

3.3 Creativity techniques

Various creativity techniques have been developed to benefit different stages in the creative process (Vernon et al., 2016). The techniques incorporated in the current training facilitated idea generation (i.e., silent brainstorming and analogical thinking) or idea evaluation (i.e., idea evaluation metric and strengths and weaknesses analysis). Each of these techniques is described in detail below.

Creativity techniques enabling *idea generation* generally employ two types of methods to facilitate the generation of ideas: (1) a stimuli-oriented method (i.e., focusing on internal or external as means to generate new ideas) and (2) a relationship-oriented method (i.e., focusing on free association or forced relationships as means to generate new ideas). Internal or external stimuli refers to the different types of stimuli used to achieve a shift in perspective with respect to the problem to foster idea generation. In techniques employing 'internal stimuli', the problem statement itself is the main stimulus to foster the generation of new ideas. In contrast, in techniques employing 'external stimuli', objects, pictures, or concepts that are unrelated to the problem statement are used to trigger new ideas to arise. The greater the perspective shift, the more likely it is that remote elements will be formed into new combinations, and hence ideas are produced that are different from each other (Dahl & Moreau, 2002).

Furthermore, free associations or forced relationships refers to the use of linking new perspectives to the problem to foster idea generation. In techniques that apply *free association*, participants follow their own train of thought and rely largely on chance and incubation. They initially generate ideas that are most accessible in memory, and, therefore, common rather than original. In contrast, *forced relationships* can be described as forcing together two or more different objects, products, or ideas to produce different kind of ideas. In forced relationships, students associate two unrelated concepts that results in forming remote associations (Daly, Christian, Yilmaz, Seifert, & Gonzalez, 2012). Remote associations are more likely to produce original ideas, as they ensure that people think outside-the-box (Isaksen, Dorval, & Treffinger, 2010).

Silent brainstorming is one of the creativity techniques in the training programme that enhances idea generation. This technique employs internal stimuli and free associations as means to come up with new ideas. In this technique, a problem statement which consists of internal stimuli (i.e., important nouns and verbs) is used to generate new ideas by means of free association with the internal stimuli. In the training, the participants were first provided with an explanation of the advantages of brainstorming individually. Specifically, they were informed that brainstorming alone and in silence is more effective than traditional brainstorming, because it allows one to generate ideas without any restrictions, guidelines or distractions (i.e., free association). This technique is focused on the quantity of produced ideas and not on the number of different conceptual categories into which the ideas can be classified (Stroebe, Nijstad, & Rietzschel, 2010). Based on the above, we hypothesized that:

> *H1a: Students who received creativity training will generate more ideas for divergent thinking tasks (i.e., fluency) than those who did not receive the training.*

Another technique that was employed in the training programmes, is *analogical thinking*. This technique employs external stimuli and forced relationships as means to come up with new ideas (Gassmann & Zeschky, 2008). In this technique, one has to associate characteristics from an external stimuli (i.e., the analogy) with characteristics of the original problem to create new ideas (Daly et al., 2012). The forced relationship means that these characteristics from an external stimuli have to be forced back to the original problem which could lead to perspective shifts. Ritter and Mostert (2016) found that the application of this technique resulted in an increase in the number of ideas that can be classified into different conceptual categories (i.e., cognitive flexibility). Based on the above, we hypothesized that:

> *H1b: Students who received creativity training will generate more different conceptual categories into which their ideas can be classified for divergent thinking tasks (i.e., flexibility) than those who did not receive the training.*

Once a large number of ideas have been generated using one or more of the idea generation techniques, it has to be decided which solution is most promising. Several techniques have been developed to enhance *idea evaluation* (Vernon et al., 2016). There are hints in the literature that techniques such as metrics to classify the feasibility and originality of an idea can be helpful in identifying how feasible and original it is (Sarkar & Chakrabarti, 2008). In addition, it has been found that a strength and weaknesses analysis enhances performance in idea evaluation accuracy (Licuanan et al., 2007).

An *idea evaluation metric* is one of the creativity techniques in the training programme that enables idea evaluation (Sarkar & Chakrabarti, 2008). In the training, participants were asked to classify all their generated ideas according to their feasibility and originality. This metric provides participants with more insights into what kind of ideas they have generated. For instance, one idea to expand businesses outside an emerging market could be "marketing". By using the metric, the idea could be put in the matrix under highly unoriginal and highly feasible. As a result, participants might realize that this idea has been mentioned a lot by other students during idea generation and, therefore, helps participants in recognizing which ideas are truly original or not. Based on the above, we hypothesized that:

> *H2a: Students who received creativity training will become more accurate in idea evaluation in terms of originality than those who did not receive the training (compared to experts).*

The *ALoU technique* is one of the other creativity techniques in the training programme. This is a strength and weaknesses analysis which stands for Advantages, Limitations, how to Overcome them and Unique qualities (Treffinger, 2007). For each idea, participants are instructed to think of its advantages, its limitations, how to overcome these limitations, and the unique qualities. The ALoU technique stresses idea evaluation as an inherently creative activity in which the implications of ideas must be explored: students are instructed to think of limitations for each idea. Next, they are asked to generate possible ways to overcome these limitations. An original idea that seems unfeasible and quite outlandish may, perhaps with some modifications, turn out to be very successful after all. To clarify this process, imagine that as a possible solution for the problem of exceeding the speed limit, the following idea is generated: "speed camera lottery pays drivers for slowing down". At first sight, a limitation of this idea would be that it is a costly idea and that it is unfair to use taxpayer money for this. A possible way to overcome this limitation is to come up with the idea to collect money from speeders to pay law-abiding drivers. In this way, the limitation that the idea is costly and involves taxpayer money is off the table. By exploring the implications of ideas and restructuring of ideas, it might help participants in recognizing which ideas are truly feasible or not. Based on the above, we hypothesized that:

> *H2b: Students who received creativity training will become more accurate in idea evaluation in terms of feasibility than those who did not receive the training (compared to experts).*

3.4 Methods

3.4.1 Participants

The current study was conducted from September 2017 to December 2017 at Maastricht University in the Netherlands. In total, 51 third-year bachelor students in international business followed the creativity training entitled 'Creative Problem-Solving for Emerging Markets'. From 51 students, 24 were female and 27 were male, and the average age was 22 (SD = 1.33), ranging from 20 to 26 years.

The research was not of medical nature, no minor or persons with disability were involved, and there were no potential risks to the participants; therefore, ethical approval was, when data collection started, not required by the Institution's guidelines and national regulations.

3.4.2 Procedure

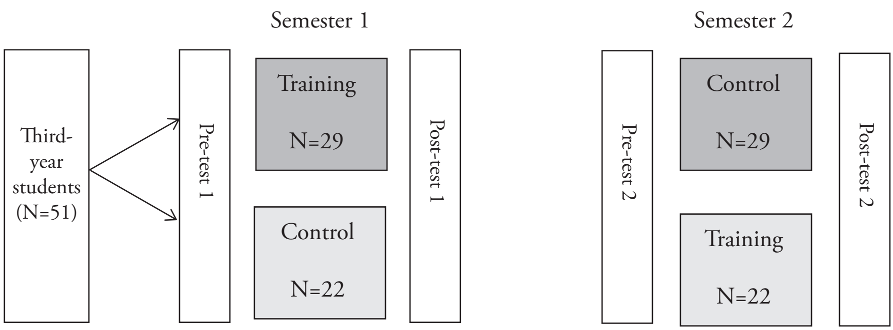
The study employed a pre-post-test within-subject design. In total, 51 undergraduates performed two treatments – a control trial and an intervention trial – which were counterbalanced and with sufficient time between trials to allow residual effects to dissipate (four weeks between trials). In the control group, students did not receive creativity training, whilst in the intervention group they received a 10-h creativity training. The same procedures were used during both semesters, which were conducted by the same experimenter and creativity trainer.

As such, 29 students attended the training in the first educational semester and then performed in the control group in the second semester, whilst 22 students performed the control group first and attended the training in the second educational semester (see Figure 3.1). As such, students participated in both experimental conditions (control and intervention), albeit at different times (within-subject design).

Students' divergent and convergent thinking skills were assessed at four time points via online surveys: at the beginning of the training programme in semester 1 (pre-test 1), at the end of the training programme in semester 1 (post-test 1), at the beginning of the training programme in semester 2 (pre-test 2), and at the end of the training programme in semester 2 (post-test 2). Each online survey took approximately 30 minutes.

At each measurement occasion, divergent thinking was measured by the often used Alternative Uses Task (AUT: Guilford, 1967) and convergent thinking by specifically designed idea evaluation tasks (see 'Measures' section). For each semester, the pre- and post-measures were the same. To prevent fatigue effects among students, the second educational semester employed equivalent versions of the tasks used in semester 1 (i.e., the versions did not differ in the types of question nor in levels of difficulty).

Figure 3.1
Experimental design: pre-post-test within-subject design



3.4.3 Creativity training

The creativity training programme is provided as a mandatory course for third-year bachelor students in business and economics at Maastricht University in the Netherlands.The creativity course (in total 10 hours) contained five weekly two-hour sessions. The course entailed an opening lecture (i.e., focus on theory) and training sessions (i.e., focuses on practice exercises in which students apply techniques on problems in the field of international business).

In the creativity training, students learned to apply four stages of creative problem solving to a wide range of problems. The four stages – problem definition, idea generation, idea evaluation, and idea implementation – are described in more detail below.

(1) Problem definition. Students received ill-defined problems in the field of international business. Ill-defined problems are often characterized by multiple goals, requiring diverse avenues of exploration that highlight a range of possible solutions (Reiter-Palmon & Illies, 2004). As such, students acquired knowledge to define the problem correctly.

(2) Idea generation. This stage is often associated with divergent thinking. Divergent thinking involves producing multiple or alternative answers to an open-ended problem (Guilford, 1959). As such, students were asked to generate different kinds of alternatives instead of focusing on one idea or solution.

(3) Idea evaluation. This stage is often associated with convergent thinking. In contrast to divergent thinking, convergent thinking can be defined as a more strongly constrained process that searches for one possible outcome for a given problem (Hommel, Colzato, Fischer, & Christoffels, 2011). In idea evaluation, ideas are checked against criteria for the task and criteria in the domain more generally, to ensure the usefulness or appropriateness of the novel ideas emerging from the idea generation stage (Lonergan, Scott, & Mumford, 2004).

(4) Idea implementation. In this stage, decisions are made based on the results of the idea evaluation stage.

Four creativity techniques were provided to facilitate the four stages of creative problem solving: silent brainstorming, analogical thinking, idea evaluation metric, and ALoU (see 'Creativity techniques' section). In the training, students are asked to solve an ill-defined problem in the field of international business. In each of the training session, they are repeatedly provided with different types of example problems, which trigger them to practice the different stages of creative problem solving. Each training session started with explanation and illustration of a creativity technique. After the explanation, students participated in a warming-up or energizer: a short group activity that is not aimed at developing creativity, but at energizing the students. The warming-up prepares the mind for the theory and training provided.

3.5 Measures

A number of measures were used to assess baseline performance and improvement in idea generation and idea evaluation. These measures were administered before and after the creativity training.

3.5.1 Idea generation

Idea generation was assessed by the Alternative Uses Task (AUT: Guilford, 1967). This divergent thinking task is widely used to evaluate creative thinking potential and the effectiveness of creativity training (Acar & Runco, 2012; Dyson et al., 2016). In this task, individuals are asked to list as many different and unusual uses for common household objects as possible in two minutes. In the first semester, these objects were a *brick* and an *umbrella*, and in the second semester a *newspaper* and a *paper clip* (both pre- and posttest). The objects were counterbalanced in each semester.

The total number of non-redundant ideas (i.e., fluency) and the total number of different conceptual categories into which ideas can be classified (i.e., flexibility) were measured for the AUT.[10] To measure flexibility, two independent raters, who were not informed of the conditions, classified each idea according to a predefined list of idea categories (see Tables A3.1 and A3.2 in the Appendix). Next, the total number of distinct idea categories is calculated. The interrater reliability of the ratings was calculated using a two-way random ICC analysis for consistency and can be considered excellent (ICC_{brick} = .92, $ICC_{umbrella}$ = .90, $ICC_{newspaper}$ = .88, $ICC_{paper\ clip}$ = .90).

3.5.2 Idea evaluation

To objectively measure participants' ability to evaluate ideas, a domain-specific idea evaluation task was developed. Prior to our experiment, 33 students from the previous cohort were asked to individually generate ideas for two similar problems: 'What can government in emerging markets do to attract new business from abroad?' and 'What can companies in emerging markets do to expand their business outside their own countries?' This resulted in respectively 285 and 227 ideas. These were further reduced to a list of 71 and 76 ideas by removing duplicates and collapsing ideas that were similar (i.e., 'improve language proficiency' and 'learn languages' were collapsed into one idea, 'learn the language').

These ideas were then rated by three experts. The experts are professionals that have worked in a private organization or government in an emerging market on similar issues as the problem in question.[11] As such, these experts had considerably more experience in the domain than students. The experts were asked to rate each idea on originality and feasibility using a Likert scale ranging from 1 (*not at all original/feasible*) to 5 (*very original/feasible*).

After their evaluation, a set of ten ideas was selected for semester 1. To prevent fatigue effects, another set of ideas was selected for semester 2 in which five ideas remained the same as in semester 1 (see Tables A3.3, A3.4, A3.5, and A3.6 in the Appendix). The sets of ideas showed sufficient high inter-rater reliability: The overall intraclass correlation coefficient (ICC, two-way random, consistency

10 The originality of ideas is not measured, because the main goal of the study was to replicate existing findings regarding the effect of creativity training on divergent thinking and the measures fluency and cognitive flexibility are considered to be sufficient. Moreover, the total number of ideas (3,748 ideas) was so large, that it would have been impossible for raters to process.

11 For instance, one expert has worked in an organization that offers support and advice to public-private partnerships in the health sector in emerging countries (i.e., Nigeria and Tanzania). Another expert has worked at the Thai government developing policies to attract new business from abroad such as tax subsidy system, visit trade shows, the establishment of an organization to help foreign investors.

analysis) was .75 for semester 1 and .74 for semester 2, and the single interrater reliability were good as well (feasibility ICC $_{\text{semester 1}}$ = .85; feasibility ICC $_{\text{semester 2}}$ = .73; originality ICC $_{\text{semester 1}}$ = .75; originality ICC $_{\text{semester 2}}$ = .80). In each measurement, students received these set of ideas for the two problems.

Students first evaluated ideas in terms of originality, i.e., the degree to which an idea is unique or novel (Putman & Paulus, 2009; Rietzschel et al., 2010). Second, students evaluated ideas in terms of feasibility, i.e., the degree to which an idea is practical or realistic (S. Kim, Chung, & Yu, 2013; Reiter-Palmon, Mumford, O'Connor Boes, & Runco, 1997). Students used a Likert scale ranging from 1 (*not at all original/feasible)* to 5 (*very original/feasible).*

Idea evaluation accuracy was measured by comparing student's evaluations with those of experts. As described earlier, experts evaluated each individual idea for originality and feasibility using the same Likert scale as was used by the students. By averaging the scores of the three experts, each idea received an *originality* and *feasibility* score. For each idea, the value of the participants' evaluations was subtracted from the average value of the expert's evaluation for that respective idea, separately for originality and feasibility. These differences were then turned into absolute differences so that larger differences reflected less accurate evaluations by the participants, regardless of direction. To easily interpret training effects, these values were reversed so that higher values indicated a more accurate evaluation. As such, idea evaluation accuracy can be seen as the degree of concordance between students and experts.

3.6 Results

The means and standard deviations for each measure at pre- and post-test for the two groups are shown in Table 3.1. Three participants did not complete the divergent thinking tasks, and, therefore, the performance of 48 participants could be analysed on the divergent thinking task.

Table 3.1
Means and standard deviations for each measure at pre-and post-test

		Idea generation (N = 48)				**Idea evaluation (N = 51)**			
		Fluency		Flexibility		Accuracy in feasibility		Accuracy in originality	
		Pretest	Post-test	Pretest	Post-test	Pretest	Post-test	Pretest	Post-test
Training	*Mean*	5.77	6.42	4.63	5.14	2.61	2.57	2.75	2.79
	SD	(2.44)	(2.61)	(1.88)	(1.78)	(0.30)	(0.29)	(0.23)	(0.21)
Control	*Mean*	5.97	6.14	4.88	5.00	2.70	2.62	2.82	2.81
	SD	(2.27)	(2.45)	(1.55)	(1.76)	(0.25)	(0.29)	(0.18)	(0.22)

A General Linear Model (GLM) for repeated measures was conducted to determine whether any change in idea generation and idea evaluation is the result of the interaction between type of treatment (i.e., intervention or control trial) and time (pre- and post-test). For each dependent variable, a separate GLM for repeated measures was conducted with treatment and time as independent variables (see Table 3.2).

Table 3.2

Test of within subject effect using General Linear Model (GLM) for repeated measures

Source	Sum of squares	df	Mean square	F value		p value	partial η^2
Test of within subject effect on idea generation: fluency							
Time	7.922	1	7.922	3.811	*	0.057	0.075
Treatment	0.083	1	0.083	0.026		0.872	0.001
Time x Treatment	2.755	1	2.755	1.344		0.252	0.028
Error	96.37	47	2.05				
Test of within subject effect on idea generation: flexibility							
Source	Sum of squares	df	Mean square	F value		p value	partial η^2
Time	4.845	1	4.845	4.105	*	0.048	0.080
Treatment	0.158	1	0.158	0.103		0.749	0.002
Time x Treatment	1.783	1	1.783	1.709		0.198	0.035
Error	49.030	47	1.043				
Test of within subject effect on idea evaluation: accuracy in feasibility							
Source	Sum of squares	df	Mean square	F value		p value	partial η^2
Time	0.193	1	0.193	9.938	***	0.003	0.166
Treatment	0.246	1	0.246	1.596		0.212	0.031
Time x Treatment	0.028	1	0.028	0.875		0.354	0.017
Error	1.571	50	0.031				
Test of within subject effect on idea evaluation: accuracy in originality							
Source	Sum of squares	df	Mean square	F value		p value	partial η^2
Time	0.019	1	0.019	0.927		0.340	0.018
Treatment	0.089	1	0.089	1.910		0.173	0.037
Time x Treatment	0.040	1	0.040	1.271		0.265	0.025
Error	1.581	50	0.032				

Note: The large effects of time can be explained by the identical tasks in the pre-and posttest (for each semester). In order to rule out learning effects (i.e., increase in fluency and flexibility due to learning on the task rather than learning due to training), we tested whether the increase in fluency and flexibility for the intervention group was larger than the increase in fluency and flexibility for the control group. As expected, we found that the intervention group reported a significant larger increase on fluency and flexibility than the control group. Both groups are expected to report increases on fluency and flexibility, but the intervention group reported a significantly larger increase due to training. Significance levels indicated as follows: * $p < .10$, ** $p < .05$, *** $p < .01$. Due to the small sample size in this Chapter, we stated the level of (marginal) significance at .10. This decision was based on Mudge, Baker, Edge, and Houlahan (2012) who argued that the observed effect size required to produce a significant result using $\alpha = 0.05$ decreases as the sample size increases. As such, p-values are dependent on sample size and studies with different sample sizes could require different observed effect sizes to yield a significant result.

3.6.1 Idea generation

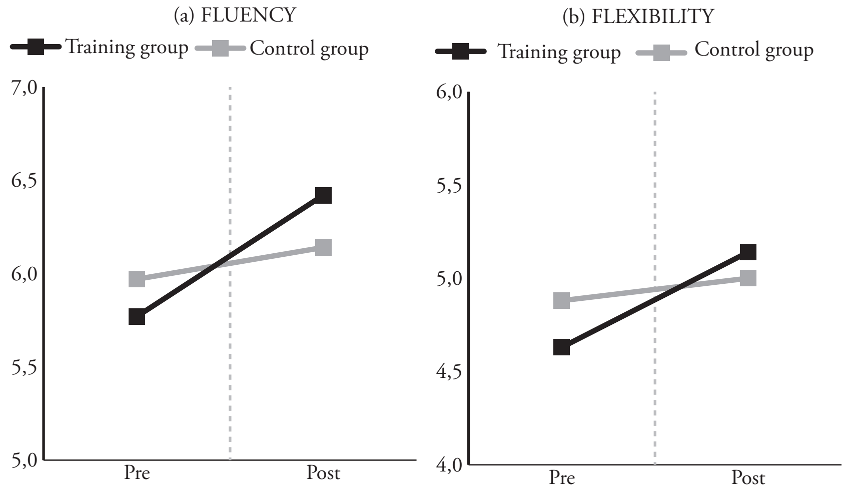
It was firstly hypothesized that creativity training would enhance the total number of generated ideas (i.e., fluency) (H1a). As shown in Table 3.1, students with training reported a higher level of fluency after training (*M* = 6.42, *SD* = 2.61) than students without training (*M* = 6.14, *SD* = 2.45). The General Linear Model (GLM) for repeated measures on fluency indicated a marginal significant main effect of time, $F(1, 47) = 3.811$, $p < .10$, partial $\eta^2 = .075$ (see Table 3.2). This indicates that students were better at post-test than pre-test. However, neither the main effect of treatment ($F(1, 47) = 0.026$, $p > .10$) nor the interaction of time with treatment were found to be significant ($F(1, 47) = 1.344$, $p > .10$). As such, hypothesis 1a was rejected.

Secondly, it was hypothesized that creativity training would enhance the number of different conceptual categories into which ideas can be classified (i.e., flexibility) (H1b). An inspection of the mean scores (Table 3.1) shows that students with training reported a higher level of flexibility after training (*M* = 5.14, *SD* = 1.78) than students without training (*M* = 5.00, *SD* = 1.76). The General Linear Model (GLM) for repeated measures on flexibility indicated a marginal significant main effect of time, $F(1, 47) = 4.105$, $p < .10$, partial $\eta^2 = .080$ (see Table 3.2). This means that students were better at post-test than pre-test. However, neither the main effect of treatment ($F(1, 47) = 0.103$, $p > .10$) nor the interaction of time with treatment were found to be significant ($F(1, 47) = 1.709$, $p > .10$). As such, hypothesis 1b was rejected.

Hence, we found that students were better in idea generation at the post-test than pre-test (see Figure 3.2). However, the interaction of time and treatment did not reach significance, and, therefore, we cannot conclude that training significantly improved idea generation skills among students.

Figure 3.2

Interactions between treatment and time on idea generation



Note: The graphs report estimates of the General Linear Model (GLM) for repeated measures. It shows the mean fluency and flexibility scores in the pre and post-test per experimental group (scale range is +½ SD and −½ SD from the average).

3.6.2 Idea evaluation

For idea evaluation accuracy, we first hypothesized that students would become more accurate in their evaluation of original ideas after training (H2a). An inspection of the mean scores (Table 3.1) shows that students without training recognized more accurately which ideas were original (M = 2.81, SD = 0.22) than students with training (M = 2.79, SD = 0.21). However, the General Linear Model (GLM) for repeated measures did not find a significant interaction of time with treatment, $F(1, 50) = 1.271$, $p > .10$ (see Table 3.2). Hence, hypothesis 2a was rejected.

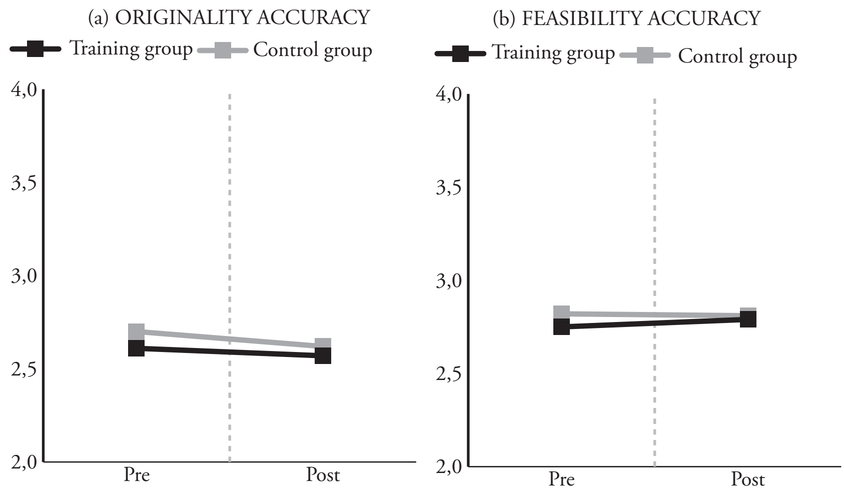
Secondly, it was hypothesized that students would become more accurate in their evaluation of feasible ideas after training (H2b). An inspection of the mean scores (Table 3.1) suggests that students without training recognized more accurately which ideas were feasible (M = 2.62, SD = 0.29) than students with training (M = 2.57, SD = 0.29). However, the General Linear Model (GLM) for repeated measures did not find a significant interaction of time with treatment, F(1, 50) = 0.875, $p > .10$ (see Table 3.2). Hence, hypothesis 2b was rejected. Interestingly, there was a strong significant

3

main effect of time, $F(1, 50) = 9.938$, $p < .01$, partial $\eta^2 = .166$ (see Table 3.2). This means that students were significantly better at post-test than pre-test.

Hence, the results suggest that creativity training did not improve idea evaluation accuracy (see Figure 3.3).

Figure 3.3
Interactions between treatment and time on idea evaluation



Note: The graphs report estimates of the General Linear Model (GLM) for repeated measures. It shows the mean originality accuracy and feasibility accuracy scores in the pre and post-test per experimental group (scale range is +½ SD and −½ SD from the average).

3.7 Discussion

3.7.1 Conclusion

There is a growing consensus that education should cultivate creative thinking skills among students to help them succeed in modern, globalised economies based on knowledge and innovation (Cachia et al., 2010; Lucas, Claxton, & Spencer, 2013). For example, in the Programme for International Student Assessment (PISA) from the OECD, students' creative thinking skills will be assessed from 2021 onwards, in addition to the existing tests in math, language, and science. This underscores the increased importance to give more attention to the development of creative thinking skills, such as idea generation and idea evaluation.

The main goal of the current study was to examine whether a cognitive-based creativity training enhances students' idea generation and idea evaluation skills. In line with previous research, the results suggest that students generated more ideas (i.e., fluency) and more different kind of ideas (i.e., flexibility) after training, but not significant. One possible explanation for this non-significant finding might be due to the small sample size (N = 48). Another possible explanation might be that post-test measures in other studies were more similar to the training context (Byrge & Hansen, 2013; Im, Hokanson, & Johnson, 2015). For example, in Im et al. (2015) participants were asked to develop lists of alternative uses for common objects during the training, and were asked to do the same in the post-test measures. Similarly, in Byrge and Hansen (2013), participants were asked to take an item from an item box (box with many different items, such as a watch, a spoon, a tissue). Next, they were asked to generate as many ideas as they could on how to improve someone else's item using their own item. Both of these creativity exercises are rather similar to the Alternative Uses Task (AUT). As such, students in previous studies were more trained to the criterion (Scott et al., 2004a).

In order to move from creativity to innovation, it is of vital importance to recognize whether the generated ideas have creative potential. Therefore, we also examined whether the training improves students' ability to recognize original and feasible ideas of others (i.e., the two components of creative ideas). In the idea evaluation tasks, participant had to rate business ideas on their originality and feasibility, and their ratings were compared with those of experts to determine accuracy. The results of this study suggest that creativity training had no effect on participants' idea evaluation performance. This finding is in line with Ritter et al. (2020) who also found no significant effect of training on participants' idea selection performance. As such, the results suggest that idea evaluation techniques, such as the idea evaluation metric and the ALoU technique do not improve idea evaluation accuracy. Research has suggested that knowledge plays a critical role in the convergent thinking process, as it is a source of ideas, and provides the knowledge necessary for assessing novelty and feasibility of ideas (Cropley, 2006). As such, a possible explanation for the fact that idea evaluation techniques do not seem to work is because these techniques are general in nature and not focused on building up knowledge and expertise.

3.7.2 Limitations and future research

Several limitations of the present study should be noted. Firstly, the creativity trainer was also the experimenter of the study. As such, the 'Clever Hans' phenomenon in which the experimenter's hypothesis is unintentionally communicated to the participants might have happened in this study (Johnson, 1911). However, this would have only affected idea generation and not idea evaluation. The experimenter's hypothesis was rather predictable with idea generation (better performance in idea generation tasks), but more difficult to define for idea evaluation since idea evaluation accuracy is highly dependent on domain knowledge (the experimenter does not know when an idea is truly feasible or original according to the experts). Nonetheless, we would still recommend for future research that the creativity trainer would be someone else than the experimenter of the study.

The second limitation relates to external validity issues. Like Ritter et al. (2020), the current study has been conducted among international business students. Even though students are often used as a research population, the population of international business students is rather specific. As such, the results cannot easily be generalized to other student populations (e.g. medicine or law). Future research should investigate the effect of creativity training on idea generation and idea evaluation on other student populations as well. Related to this, students have only been asked to generate ideas for one divergent thinking task (i.e., the Alternative Uses Task). As such, it is unknown whether the training effects can be generalised to other divergent thinking tasks (e.g. consequences task). Even though the main contribution of the current study was to investigate the effect of creativity training on convergent thinking, future research should use multiple measures of divergent thinking to ensure the external validity of the results.

Furthermore, the length of training is positively related to creativity training effectiveness (e.g. Mathisen & Bronnick, 2009; Scott et al. 2004). Mathisen and Bronnick (2009) found, for example, evidence for the effectiveness of both a 1-day and a 5-day training course. Still, they found their 5-day training to induce stronger effects. The current creativity training consists of only 10 hours of explicit training and this might be one of the reasons why the training appears to be unsuccessful. However, a recent study evaluated a 140 hours training on convergent thinking and also found no improvements in convergent thinking (Ritter et al. 2020). As such, simply increasing the length of a training may not be sufficient to enhance performance in convergent thinking. Next to training, there are a number of other approaches that have been proven to be successful in enhancing idea evaluation skills, such as task familiarity, mood, personality, regulatory focus, culture, and gender (e.g.

Berg, 2016; Förster, 2009; Herman & Reiter-Palmon, 2011; Kaufman, Niu, Sexton, & Cole, 2010; McCarthy, Chen, & McNamee, 2018). For instance, Basadur et al. (2000) found that people who had a natural tendency to avoid premature convergence - a trait related to openness – performed better in idea evaluation tasks than people without this tendency. As such, future research, dedicated to evaluation skills, should take a step further in testing training programmes with other theoretical foundations, such as a personality or social interactional model instead of a purely cognitive-based training.

Finally, this study has conducted General Linear Models (GLM) for repeated measures as a means to compare students who received the creativity training (intervention group) with those who did not (control group). The repeated measures nature of this study allowed us to compare the intervention group in semester 1 with a control group. However, since our training was offered to all students (for ethical reasons), there was no suitable comparison group that could participate in the long-term follow-up (i.e., semester 2). Therefore, the students who received the training in the second semester were compared with the students who had received the training earlier during the first semester. This might have been a reason why no effect of the creativity training on idea generation (divergent thinking) and idea evaluation (convergent thinking) could be found. However, the idea behind this was that the effect of training would be stronger visible immediately after the training, and would diminish after four weeks. We indeed found that the intervention group in semester 1 (who served as control group in semester 2) reported lower levels of fluency (M = 5.95, SD = 2.56) and flexibility (M = 4.79, SD = 1.81) after four weeks of no training than immediately after training (fluency: M = 6.21, SD = 2.59; flexibility: M = 4.97, SD = 1.78). This decline in creativity scores is in line with a recent study that indicated that four weeks may be sufficient to let training effects diminish (Meinel, Wagner, Baccarella, & Voigt, 2019). As such, it could be very well that the participants who attended the training in the first semester were indeed a valid control group for the students receiving the training in the second semester. Nonetheless, we would still recommend for future research to add an extra control group consisting of a random sample of comparable students who have not attended the training.

3.7.3 Practical implications

The current study has provided insight into how a cognitive-based creativity training is not successful in strengthening idea evaluation skills. This is an important insight, because many cognitive-based training programmes assume that convergent thinking skills, such as idea evaluation, are naturally developed in the training programmes.

However, in line with previous research, we show that convergent thinking skills are not naturally developed within the context of a cognitive-based training.

As such, our key contribution is showing that a cognitive-based creativity training does not affect the evaluation accuracy of novel and feasible ideas, and, therefore, idea evaluation seems to be a more complex process to enhance than idea generation. However, spotting the novelty and feasibility is a crucial step in getting people to pay attention to the ideas already generated. Without it, ideas will not be captured and developed to add value in solving problems. Institutions who would like to or already have implemented cognitive-based creativity training should be aware of that such a training does not automatically foster convergent thinking among students. As indicated earlier, a number of other approaches have been proven to be successful in enhancing idea evaluation skills, such as task familiarity, mood, personality, regulatory focus, culture, and gender (e.g. Berg, 2016; Förster, 2009; Herman & Reiter-Palmon, 2011; Kaufman et al., 2010; McCarthy et al., 2018). As such, it might be that training programmes with other theoretical foundations, such as personality or social interactional model may be more successful in enhancing convergent thinking skills.

3.8 Appendix

Table A3.1

Idea categories Alternate Uses Tasks: umbrella and brick

Umbrella (used for protection against the rain)	*Brick (used for building walls and houses)*
As decoration	As decoration
As a weapon or to defend yourself	As a weapon or to defend yourself
For (other) weather conditions	To throw
As a bowl or bag (e.g. basket)	To break things
To hide someone of something	To support
As a tool	As a tool
To draw, open and push things	Construction (other than buildings)
Reaching unreachable spots (e.g. scratching)	To draw or scratch
To balance	As a weight
As clothing	As clothing
To cover things in the house or garden	As other objects (e.g. cooking objects)
As a walking stick	As a barrier/ to block
As a pointing stick	As a BBQ, grill or fireplace
Flying-related usage	As a plant pot
Sport-related usage	Sport-related usage
Photoshoot	As a landmark
As a toy/game	As a toy/game
As furniture	As furniture

Table A3.2

Idea categories Alternate Uses Task: paper clip and newspaper

Paper clip (used for holding pieces of paper together)	*Newspaper (used for reading)*
As decoration	As decoration
As a weapon	As a weapon
As a scratching or carving device	To write or draw on
As a needle or perforator (e.g. to poke someone)	To set in fire
As a hook or hanger	As a material to create something
As a toy	As a toy
As clothing or hair accessories	As clothing
To repair something (e.g. clothes)	To soak, dry or clean
For keeping things together or connecting things (not as fixing things and not in hair)	Holding things or fill up
Electricity-related usage	To wrap something (e.g. food)
As a cleaning tool	To be undercover (e.g. hiding)
As an opening tool	As insulation material
Other types of tools (not in any of the above)	As a fan
	To roll-up

Table A3.3

Semester 1.

Ideas for 'what can government in emerging markets do to attract new business from abroad?'

Feasibility	*High*	• Visit trade fairs	• Attracting talents early on (e.g. during studies) • Ease visa issues for foreign employees
	Low	• Creating favourable policy regulation (e.g. tax breaks/free-export zones) • Good infrastructure for business (e.g. reliable water and electricity connection • Reduce corruption • Reduce bureaucratic procedures • Policy focuses on foreign business	• Guarantee access to international currency • Refrain from having a radically religious cultures
		Low	*High*
		Originality	

Table A3.4

Semester 1.

Ideas for 'what can companies in emerging markets do to expand their business outside their own countries?'

Feasibility	*High*	• Join start up communities • Hire experts from countries they have interest in • Start exporting before starting an own factory/office abroad • Marketing • Send employees abroad for conferences or other forms of training • Have local partners to facilitate the market entry	
	Low	• Partner with a foreign firm • Invest in a project that will happen in a foreign country • Quality competition	• Go to more corrupt countries and bribe government to allow easy set up
		Low	*High*
		Originality	

Table A3.5

Semester 2.

Ideas for 'what can government in emerging markets do to attract new business from abroad?'

Feasibility	*High*	• Set up business parks to attract business	• Provide services that support foreign business (e.g. personal and work related support such as assistance for expats in housing/ schooling and administrative issues) • Identifying their own local competitive advantage and exploit it (e.g. Southeast Asia van serve as a logistical hub between Europe/Japan/rest of Asia and Australia)
	Low	• Creating favourable policy regulation (e.g. tax breaks/free-export zones) • Good infrastructure for business (e.g. reliable water and electricity connection • Reduce corruption • Reduce bureaucratic procedures • Loosen ethical standards (e.g. be more willing to accept bribes from foreign firms) • Encourage cooperation between companies and governments (e.g. offer partnership programs) • Protect workforce (e.g. safety for employees)	• Motivate local business to cooperate or do joint venture • Give population more liquidity in order to stimulate the own economy in order to attract investments from outside • Promote the country's assets and/ or its natural resources abroad • Meet and greet between politics and business • Ease visa issues for foreign employees • llustrate possible demand in country • Host fairs or networking events where governments, business and entrepreneurs can meet each other and promote themselves
		Low	*High*
		Originality	

Table A3.6

Semester 2.

Ideas for 'what can companies in emerging markets do to expand their business outside their own countries?'

Feasibility	*High*	• Go to more corrupt countries and bribe government to allow easy set up • Innovation in creating new products • Product competition	• Learn from others and copy their ideas • Outsourcing/offshoring parts of their business
	Low	• Partner with a foreign firm • Adapt their own business to foreign markets • Get bought by Google or some other large company • Quality competition	• Hire experts from countries they have interest in • Marry someone from the other country to get insights into the culture • Use their cultural background as advantage (e.g. shoes using leather of indigenous people) • Creative competition (e.g. cost leadership or differentiation) • Marketing • Price competition (i.e., sell at lower cost) • Sell human capital
		Low	*High*
		Originality	

4

Creative idea forecasting: The effect of task exposure on idea evaluation

Abstract

In various settings, people are responsible for evaluating ideas generated by others while they were not involved in the idea generation process, and, as such not exposed to the task. However, little is known on how this lack of task exposure affects creative forecasting. This study therefore examines the effect of task exposure on creative idea evaluation. For this purpose, 1864 German students evaluated ideas on their creativity, originality and feasibility. Their ratings were compared to content experts' and creativity experts' ratings. The students were randomly assigned to one of the following conditions: task exposure (i.e., in this condition they had to generate and evaluate ideas for the same task) or no task exposure (i.e., in this condition they had to generate ideas for a different task than the idea evaluation task). The results showed that task exposure improves students' ability to accurately recognize creative, original and highly feasible ideas. Together, the results highlight the importance to carefully reconsider whether or not people should be exposed to the task before evaluating other's ideas.

This chapter is joint work with Barbara Belfi, Lex Borghans, and Philipp Seegers, and is under revise and resubmit at Psychology of Aesthetics, Creativity, and the Arts. We thank the participants of the 2019 European Collaborative Creativity Conference (EC3) and Creativity Conference Southern Oregon University (SOU) for their valuable comments and suggestions on earlier versions of this paper.

4.1 Introduction

Over the past few decades, there have been several cases where people have failed to recognize creative ideas generated by others (Elsbach & Kramer, 2003; Mainemelis, 2010). For example, J.K. Rowling's original version of Harry Potter was rejected at first by publishers but eventually became one of the all-time best-selling fiction books; the Kentucky Fried Chicken (KFC) recipe met with countless rejections by restaurant owners before it was accepted; the initially rejected idea of large electrostatic displays were eventually integrated into more than half of Hewlett-Packard's instruments (Nemeth, 1997).

In real-life settings, both at work as in private life, people who were not involved in the generation of a certain idea are burdened with the task to evaluate the potential of that idea (e.g. Berg, 2016). In education, for example, parents and teachers often evaluate students' ideas without any exposure to the task or problem at hand. To date, little is known how this lack of task exposure impacts creative forecasting, in which ideas that are potentially feasible as well as novel have to be detected (Corazza, 2016). Hence, the main aim of the present research is to investigate the effect of task exposure on the evaluation of potentially creative ideas generated by others.

Much of the prior research on creative forecasting has focused on assessing people's accuracy in predicting original features of new ideas (e.g. Basadur, Runco, & Vega, 2000; Grohman, Wodniecka, & Kłusak, 2006; Licuanan, Dailey, & Mumford, 2007; Runco & Smith, 1992; Silvia, 2008; Zhu, Ritter, & Dijksterhuis, 2019). More concretely, these studies have investigated whether idea evaluation by the people who have generated the ideas (i.e., *intrapersonal* evaluation) is more accurate than idea evaluation done by people who have not generated the ideas (i.e., *interpersonal* evaluation). These studies have generally come to the conclusion that people are more accurate in evaluating their *own* ideas than *other's* ideas (Grohman et al., 2006; Runco & Smith, 1992; Silvia, 2008).

There are two main reasons why people are better at intrapersonal evaluation than interpersonal evaluation (see Runco & Smith, 1992; Zhu, Ritter, & Dijksterhuis, 2019). Firstly, a person who generated ideas will know the associative history of each idea, and know what motivated particular directions of thought. Secondly, he or she will know what other ideas were rejected in favor of those that were kept.

However, a drawback of intrapersonal evaluation is that people may be sensitive to ownership bias. Ownership bias is a decision-making bias that leads to a tendency to prefer one's own ideas over others (Beggan, 1992; Beggan & Brown, 1994; Nikander, Liikkanen, & Laakso, 2014). According to ownership bias, a person who has previously generated an idea and has to rate that exact same idea will favor that idea. This bias might come to the surface in real-life settings where people have to judge their own ideas (intrapersonal evaluation). However, real-life settings are dominated by interpersonal evaluation, in which people have to judge ideas that were created by others without any task exposure (Berg, 2016).

Nevertheless, we argue that this lack of task exposure may not be optimal for recognizing creative ideas. Another way to prevent ownership bias, but still reap the beneficial effects of intrapersonal evaluation, would be to let people generate ideas for a task themselves, before they have to evaluate others' ideas. This task exposure may be key to enhance interpersonal accuracy in creative forecasting for several reasons. Firstly, people with task exposure will be more cognizant of the difficulty of the problem, question, or task at hand, as well as the specific circumstances (Runco & Smith, 1992). Most importantly, people with task exposure have previously generated ideas for the task and, as such, they are asked to rate ideas with several comparison points (Goldstone, Medin, Halberstadt, & Cognition, 1997; Hyypiä & Parjanen, 2013; Kunreuther, Novemsky, & Kahneman, 2001; Tesauro, 1989). In contrast, people without task exposure cannot compare the ideas with their own generated ideas, and, as such, have no comparison points. This is in line with Tversky's (1977) *diagnosticity effect*, in which a comparison between two alternatives to a target is influenced by having a third alternative at hand. According to Tversky's contrast model, the context of a comparison may influence the weights assigned to the features of an object's representation. As such, people with task exposure have more information and thus more comparison points regarding the quality of other ideas while people without task exposure do not have that knowledge. The exposed group should thus produce more accurate ratings – on all measures. More specifically, Kaufman et al. (2008) found that people with more domain knowledge are better able to recognize the creativity of ideas than people with no domain knowledge. Likewise, Denker et al. (2016) found that people with higher levels of domain knowledge are generally better able to recognize the originality and feasibility of ideas. Even though domain knowledge is not exactly the same as task exposure, it can be argued that task exposure is related to domain knowledge since it provides more information and comparison points regarding the quality of ideas. As such, we hypothesized the following:

> H1. Compared to experts, students with task exposure will rate ideas more accurate on (a) creativity, (b) originality and (c) feasibility than students without task exposure.

Furthermore, according to the *search for ideas in associative memory* (SIAM) model, ideas are generated by combining existing knowledge, by forming new associations, or by applying knowledge to a new domain (Nijstad & Stroebe, 2006). This model predicts that the majority of ideas will be semantically related to one another, and, therefore, it may be easier for people to generate more common and feasible ideas than truly novel ideas (Nijstad & Stroebe, 2006; Rietzschel, Nijstad, & Stroebe, 2007). According to the SIAM model, people with task exposure will activate a search for ideas in their associative memory. This associative memory is likely to give them more comparison points to highly feasible ideas. In contrast, people without task exposure have not (yet) activated their associative memory, and, therefore, have no immediate access to highly feasible ideas. Based on the SIAM model, we hypothesized the following:

> H2a. Compared to experts, students with task exposure will be more accurate in discriminating highly feasible ideas from less feasible ones than students without task exposure.

With regard to the evaluation of ideas in terms of their creativity and originality, research has further showed that highly creative and original ideas are relatively rare (Huber, 1998; Sharma, 1999). However, expertise with a certain topic may make creative and original ideas on that topic less rare, because people could have been exposed to creative and original ideas (Kaufman, Cropley, Baer, Reiter-Palmon, & Sinnett, 2013). Therefore, expertise could make an idea less creative or original because it has already been seen, and therefore may not be viewed as novel or creative anymore. Generally, experts have spent years either generating or evaluating ideas in their domain, comparable to the 10 years of deliberate practice needed to make a substantial contribution to a field (Ericsson, Roring, & Nandagopal, 2007). In those years, they have been exposed to numerous ideas, and this could generally make them view ideas as being less creative and less original. That being said, highly creative and original ideas are relatively rare, and if a truly novel idea comes along, an expert will be better able to recognize this highly novel idea due to his or her expertise. In a similar vein, task exposure may make people perceive creative and original ideas as less creative and original, because people could have seen creative and original ideas passing by (in their thinking processes). If a truly novel or creative ideas comes along, people with task exposure may be better able to recognize this highly novel or creative idea. As such, in line with experts, people with task exposure may become better able to recognize the (truly) highly creative and original ideas than lay people without any exposure to the task. Hence, we hypothesized the following:

> H2b. Compared to experts, students with task exposure will be more accurate in discriminating (truly) highly creative and original ideas from less creative and original ones than students without task exposure.

4.2 Methods

4.2.1 Sample and procedure

A sample of 1864 German undergraduate and graduate students (1155 women and 709 men), aged 17 to 36 (M = 22.90, SD = 3.23) took part in the present study (see Table 1). The study was part of the 12th round of the online survey research project *"Fachkraft 2030"*. The first data collection took place in September 2012. Since then, new data collections have taken place every six months in March and September (see Seegers, Bergerhoff, Hartmann, & Knappe, 2019).

Students were invited to participate in the survey via the database of *Jobmensa*, which is the largest network for student jobs and internships in Germany. Students received an invitation and up to two reminders per email to participate. It took them approximately 30 minutes to complete the survey. Participation was incentivized by a total prize money of €1,950 in Amazon vouchers: one voucher of €500 and 29 vouchers of €50.

The survey consisted of questions about students' present *study experiences*, current participation in *student jobs* and *future career expectations*. At the end of the questionnaire, students were asked to participate in a variety of psychological tests (e.g. personality and IQ tests). For the purpose of the current experiment, a set of different *creativity tasks* were added to these psychological tests. In total, 2485 students completed this part of the survey.

The *creativity tasks* consisted of five divergent thinking tasks: three function-first tasks (Baas, De Dreu, & Nijstad, 2011; de Buisonjé, Ritter, de Bruin, ter Horst, & Meeldijk, 2017; Ritter, van Baaren, & Dijksterhuis, 2012) and two form-first tasks (Alternate Uses Task, Guilford, 1967).[12] These different tasks were randomly

12 *Function-first* problems state the desired outcome, e.g. "How to transport baked beans?" for which solutions are sought. Form-first problems, in contrast, state the solution, e.g. "A Tin Can" and seek different possible uses (see Cropley, Cropley & Sandwith, 2017).

distributed over the questionnaire. As such, each student was randomly asked to generate ideas for one of these five divergent thinking tasks.

For the function-first tasks, students were asked to generate as many ideas as they could for one of the following problems: 'How to improve the use of public trains', 'How can you make teaching in your university better' or 'How to make the waiting time at cash desks more bearable'. For the form-first tasks, students were asked to generate as many different and unusual uses for a 'newspaper' or a 'brick' (Alternate Uses Task, AUT: Guilford, 1967). To ensure that students were actually engaged and exposed to idea generation, 621 students that spent less than 1 minute on the task were excluded from analyses (24.99%). This decision is based on Benedek, Mühlmann, Jauk, and Neubauer (2013) finding that a minimum of 1 minute on the task is needed to reliably assess creativity outcomes (e.g. fluency and originality).

In the *task exposure* condition (N=783), students first generated ideas for 'How to improve the use of public trains' and then evaluated others' ideas for this *same* task. In contrast, in the *no task exposure* condition (N = 1081), students first generated ideas for a *different* task before evaluating others' ideas for the train task.[13,14]

In the evaluation of ideas, students were asked to rate four ideas on how to improve the use of public trains (see Materials). They were instructed to carefully read these ideas and to evaluate each idea on their creativity, originality and feasibility (Dean, Hender, Rodgers, & Santanen, 2006). Creativity is the degree to which the idea is both feasible and original. Originality is the extent to which the idea is novel, out of the ordinary. Feasibility is the ease with which the idea can be implemented (given the current context such as financial resources, infrastructure, time required, legal issues).

13 Students in the no task exposure condition generated ideas for 'How can you make teaching in your university better', 'How to make the waiting time at cash desks more bearable' or alternate uses for a 'newspaper' or 'brick'.

14 It might be argued that the nature of the task (function-first versus form-first) might impact idea evaluation. Therefore, we tested whether students who generated ideas for a function-first task evaluated ideas differently than students who generated ideas for a form-first task. The results showed that there were no significant differences in the idea evaluation in terms of creativity, originality and feasibility between students who first generated ideas for a function-first task and students who first generated ideas for a form-first task.

Table 4.1

Descriptive statistics

	N	*Mean*	*SD*	*Min*	*Max*
Background information					
Gender (ref.=female)	1864	0.38	0.49	0	1
Age (in years)	1855	22.90	3.23	17	36
Migration background (ref.=German)	1561	0.27	0.45	0	1
Parental educational level					
No degree	1864	0.07	0.25	0	1
Low degree	1864	0.29	0.45	0	1
Middle degree	1864	0.16	0.37	0	1
High degree	1864	0.44	0.50	0	1
Educational information					
University	1771	0.63	0.48	0	1
University of applied sciences	1771	0.37	0.48	0	1
Academic year	1864	2.68	2.01	0	10.50
Field-of-study					
Art/Music	1864	0.02	0.15	0	1
Business/Economics	1864	0.15	0.36	0	1
Engineering	1864	0.21	0.41	0	1
Language/Culture	1864	0.16	0.37	0	1
Medicine/Health	1864	0.04	0.21	0	1
Natural Sciences	1864	0.21	0.41	0	1
Social Sciences	1864	0.20	0.40	0	1

Note: The "Fachkraft-2030" sample shows no systematic deviations from the German student population. The sample compares well to the population data on age, type of higher education institution, state of study and field of study that are available at the German Statistics Agency (Seegers et al. 2019). In terms of gender, the sample shows an overrepresentation of female students. However, this overrepresentation is also present in comparable student surveys like the government funded Sozialerhebung, which is collected systematically on the university level. Given that the "Fachkraft-2030" sample is collected in cooperation with Studitemps / Jobmensa one could assume that working students might be overrepresented in the "Fachkraft-2030" sample. When compared to the Sozialerhebung no such bias can be found (Seegers et al. 2019).

4.2.2 Materials

To objectively measure students' ability to evaluate ideas, we used a validated measurement tool by de Buisonjé et al. (2017). In this measurement tool, 72 ideas on 'how to improve the use of public trains' were rated by ten creativity experts (e.g. creativity researchers, social behavioral researchers, art-academy teachers). According to the Consensual Assessment Technique (CAT: Amabile, 1982), these creativity experts rated independently each idea on creativity using a Likert scale ranging from 1 (not at all creative) to 5 (very creative). To measure students' ability to evaluate ideas in terms of originality and feasibility, the experts additionally rated each idea on originality and feasibility using a Likert scale ranging from 1 (not at all original/feasible) to 5 (very original/feasible).

For this study, the original list of 72 ideas was reduced to a more balanced list of 39 ideas that were evenly distributed in terms of originality and feasibility (see Online Supplementary Material). The overall intraclass correlation coefficient (ICC, two-way random, consistency analysis) was .88, and also the single interrater reliabilities were excellent (creativity ICC = .85; originality ICC = .90; feasibility ICC = .88). Subsequently, the ideas were also rated by fourteen content experts who have worked on similar issues in a national railway company (e.g. a rail systems engineer, IT architect, policy advisors).[15] By averaging the scores of the experts, each single idea received a creativity, originality and feasibility score, separately for content experts and creativity experts.

From the pool of 39 ideas, we randomly selected four ideas for each participant to rate. Because participants evaluated four different ideas, each participant ended up rating a mix of relatively novel and relatively conventional ideas.

4.2.3 Idea evaluation accuracy

Two indexes of idea evaluation accuracy were used to measure two relevant aspects of the evaluation of ideas: (a) overall level of accuracy and (b) the ability to discriminate (i.e., to recognize highly feasible, original and creative ideas).

The overall level of accuracy was used to provide insight on the degree to which students' average rating on the four ideas corresponded with those of the experts' average ratings on the four ideas, separately for each dimension of idea evaluation. As such, this measure indicated whether students' averagely over- or underestimated ideas compared to experts' ratings (independent of idea). For instance, people may have a tendency to see things in a positive light, and therefore to generally overestimate all ideas on their feasibility.[16]

The ability to discriminate (i.e., to recognize highly feasible, original and creative ideas) was used to provide insight in the degree to which students' ratings on each

15 The experts have worked two or more years in producing and evaluating ideas for similar problems at the national railway company (14% of the experts with two years of experience, 14% of the experts with three years of experience, and 71% of the experts with more than four years of experience).

16 A student might rate all four ideas with the score 'three' on a Likert scale ranging from 1 (not at all feasible) to 5 (very feasible). Therefore, on average, the score will be 'three'.

idea corresponded with those of the experts' ratings on each idea, separately for each dimension of idea evaluation. As such, this measure indicated whether students' ratings for the high and low quality ideas were in line with experts' ratings. For example, a student who received four ideas from which two were indicated as low quality ideas (low feasibility) and two were high quality ideas (high feasibility) by the experts, may rate all, or some of these ideas in the same way as the experts did. [17]

4.2.4 Statistical analysis

To analyze students' overall level of accuracy in evaluation, we estimated the following regression:

$$y_{ij}=\beta_0+\beta_1\, idea_{ij}+\beta_2 order_{ij}+\beta_3\, reference\ point_{ij}+\beta_4\, ownership_{ij}+\beta_5\, Task\ exposure_i+\varepsilon_{ij}$$

where i denoted individuals (i.e., students) and j denoted the four ideas each student has to evaluate. Y_i indicated students' mean value of the four evaluated ideas, separately for creativity, originality and feasibility. From a total set of 39 ideas, each student randomly received four ideas. To control for the fact that each student rated a different set of ideas, we included idea fixed effects ($idea_{ij}$). In addition, to control for the fact that each student rated a different order of ideas, we included idea order fixed effects ($order_{ij}$). In this setting, each idea in the set of four ideas will be compared to the other three ideas in the set. As such, one might perceive an idea as more original when it is accompanied by three rather common ideas. However, the same idea will be perceived as less original when the other three ideas are relatively more original. Therefore, students may differ in their evaluations of the same idea. We therefore included the average score of a student on the three other ideas that he or she evaluated ($reference\ point_{ij}$). Moreover, students who have previously generated an idea and have to rate that exact same idea may favor that idea.[18] Therefore, we control

---

17 A student might rate the low quality ideas with the score 'one' while the high quality ideas are rated with the score 'five' using a Likert scale ranging from 1 (not at all feasible) to 5 (very feasible). Likewise as in the other example, the student's average score will be 'three'. Therefore, it is vital to not only assess the overall level of accuracy, but also the ability to discriminate between ideas.

18 In the task exposure condition, due to the randomization of ideas in the evaluation task, some students had to evaluate similar ideas than they had previously generated themselves. By means of text mining and text search analysis, we found out that 4.5% of the students indeed had to evaluate one or more ideas that they had earlier generated themselves, while 95.5% of the students never had to evaluate an idea that they had previously generated themselves. More particularly from the 39 ideas in the idea evaluation task, 'Free WIFI in the train', 'Create faster trains', and 'Make train travel cheaper of for free' were most often generated by students. Moreover, we investigated whether there was an ownership effect. The results of this

for this potential 'ownership effect' by including a control variable (*ownership*$_{ij}$). This control variable is a binary variable that takes value 1 for students who generated and evaluated the same idea, and zero for students who generated different ideas as the evaluated ideas. The variable of interest, *Task exposure*$_i$, is a binary variable that takes value 1 for students who have been priory exposed to the same idea generation task, and zero for students who were not exposed to the task. Then, the student's estimated mean value was compared to the content experts' and creativity experts' mean value (see Figure 4.1). Thus, the difference between students' average rating and experts' average rating indicated the overall level of accuracy: whether students' overall evaluation is less or more in line with experts' evaluation.

To analyze students' ability to discriminate, we estimated the following regression:

$$y_{ij}=\beta_0+\beta_1 \textit{reference point}_{ij}+\beta_2\ \textit{ownership}_{ij}+\beta_3\ \textit{expert mean each idea (centered)}_j\ + \beta_4 \textit{Task exposure}_i+\beta_5 \textit{expert mean each idea (centered)}_j{*}\textit{Task exposure}_i+\varepsilon_{ij}$$

where i denoted individuals (i.e., students) and j denoted the four ideas each student has to evaluate. Y_{ij} indicated students' value for each idea j, separately for creativity, originality and feasibility. Likewise as above, each idea in the set of four ideas will be compared to the other three ideas in the set (*reference point*$_{ij}$), and we control for a potential ownership effect if individual i also generate an idea similar to j (*ownership*$_{ij}$). Moreover, we included the content experts' ratings as a centered mean for each idea (*expert mean each idea (centered)*$_j$). Similarly as above, *Task exposure*$_i$, is a binary variable that takes value 1 for students who have been priory exposed to the same idea generation task, and zero for students who were not exposed to the task. The variable of interest, *expert mean each idea (centered)*$_j$***Task exposure*$_i$, is an interaction term of the students' ratings with the content experts' rating, separately for each idea. In this way, we can investigate whether the ratings of high and low quality ideas of the students in the task exposure condition differ significantly from the students in the no task exposure condition. In Figure 2, the content experts' ratings are put on

examination revealed that there indeed were ownership effects for feasibility ratings (Cohen's d = .49, $p < .01$), wherein students rated their own ideas significantly higher on feasibility (M = 4.03, SD = 1.29) than students who rated others' ideas (M = 3.38, SD = 1.37). In contrast, the results revealed a negative relationship between the ownership variable and creativity ratings (Cohen's d = .38, $p < .05$). This means that students rated their own ideas lower on creativity (M = 2.59, SD = 1.19) than students who rated others' ideas (M = 3.06, SD = 1.29). Lastly, no relationship was found between the ownership variable and originality ratings (Cohen's d = .04, $p > .05$).

the diagonal grey line for each idea. Thus, a negative slope means that there is more discrepancy between students' and experts' ratings for the lower and higher quality ideas, while a positive slope means that there is less discrepancy between students' and experts' ratings for the lower and higher quality ideas (see Figure 2). In other words, students' ratings are less or more in line with experts' ratings for the lower and higher ideas (negative respectively positive slope).

4.3 Results

Concerning hypothesis 1, we expected that students with task exposure would rate ideas more accurate on creativity, originality and feasibility (compared to experts). The linear fixed-effect model shows a significant negative difference between students with task exposure and students without task exposure for creativity (b = −0.06, $p < .05$, d = .11; see Table 2). Also for originality, the linear fixed-effect model shows a significant negative difference between students with task exposure and students without task exposure (b = −0.07, $p < .05$, d = .14; see Table 2). This means that students who were first exposed to the task by generating ideas, rated ideas significantly lower on both their creativity and originality than students who were not exposed to the task. Figure 1 shows that the ratings of students with task exposure were more in line with those of experts than the ratings of students without task exposure (both content experts and creativity experts).

Contrasting to our predictions, the linear fixed-effect model shows a significant positive difference for feasibility between students with task exposure and students without task exposure (b = 0.06, $p < .05$, d = .09; see Table 2). This means that students who were first exposed to the task by generating ideas, rated ideas significantly higher on feasibility than students who were not exposed to the task. Figure 1 shows that the ratings of students without task exposure were more in line with those of content experts than the ratings of students with task exposure. Hence, hypothesis 1 was only supported for creativity and originality and rejected for feasibility.

Table 4.2

Fixed effects predicting creativity, originality and feasibility ratings

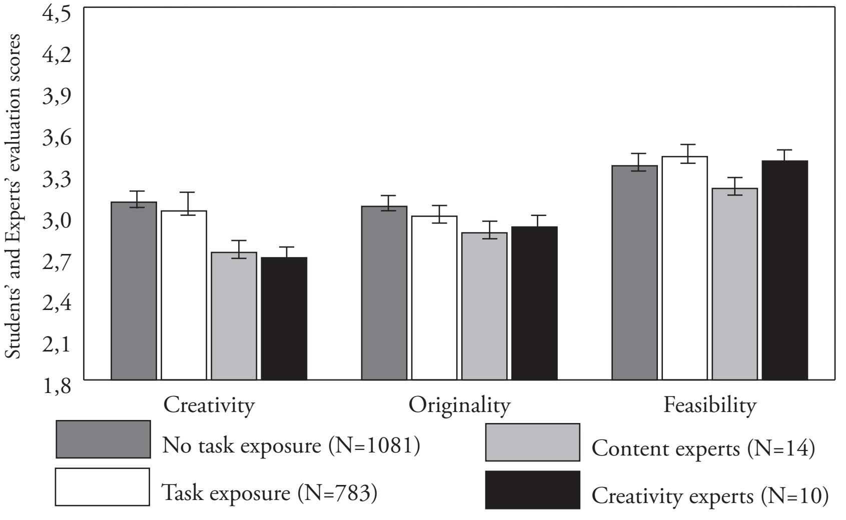
	Creativity					Originality					Feasibility				
	b	*SE*	*t*	*p*	*Effect size*	*b*	*SE*	*t*	*p*	*Effect size*	*b*	*SE*	*t*	*p*	*Effect size*
Intercept	5.40	(0.16)	34.95	<.001 ***		4.93	(0.18)	27.85	<.001 ***		4.55	(0.13)	33.92	<.001 ***	
Reference point other ideas	-0.48	(0.04)	-11.46	<.001 ***		-0.44	(0.05)	-8.86	<.001 ***		-0.26	(0.03)	-8.36	<.001 ***	
Ownership bias	-0.08	(0.19)	-0.42	.678		0.30	(0.20)	1.48	.137		0.08	(0.19)	0.43	.670	
Task exposure	-0.06	(0.03)	-2.03	.043 *	0.11	-0.07	(0.03)	-2.23	.026 *	0.14	0.06	(0.03)	2.32	.021 *	0.09
Number of unique observations	1849					1858					1859				
Number of observations	7396					7432					7433				
R-squared	0.20					0.15					0.31				

Note. Standard errors in parentheses. Significance levels indicated as follows: * $p < .05$, ** $p < .01$, *** $p < .001$.

4

Figure 4.1

Students' and experts' evaluation scores in three dimensions of idea evaluation



Note: The graph reports estimates of linear fixed-effects models (on idea level). It shows the mean evaluation scores for four evaluated ideas per criteria (scale range is +1 and -1 SD from the average). Error bars represent the confidence interval from the linear fixed-effect models.

Furthermore, hypothesis 2a predicted that student's ability to discriminate highly feasible ideas from less feasible ideas would be more accurate among students with task exposure than students without task exposure. As shown in Table 3, this is supported, as the main effect (b = 0.07, p < .05) and the interaction term show a significant positive effect (b = 0.08, p < .05). This means that students who were first exposed to the task by generating ideas became better in recognizing the highly feasible ideas (see Figure 2). Hence, we found support for hypothesis 2a.

Finally, hypothesis 2b predicted that students' ability to discriminate highly creative and original ideas from less creative and original ideas would be more accurate among students with task exposure than students without task exposure. However, Table 3 revealed that the interaction term in both the models for creativity (b = 0.03, p > .05) as well as originality (b = 0.05, p > .05) do not show a significant difference between students with task exposure and students without task exposure. This means that students who were first exposed to the task by generating ideas did not become better in recognizing the highly creative or original ideas (see Figure 2). Hence, we found no support for hypothesis 2b.

Table 4.3
Ordinary Least Squares (OLS) regression predicting creativity, originality and feasibility ratings

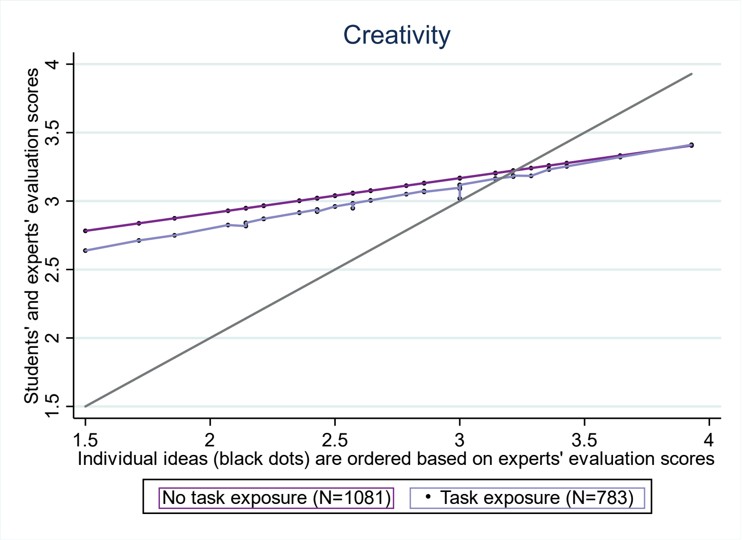
	Creativity				Originality				Feasibility			
	b	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*
Intercept	4.67	(0.14)	32.59	<.001 ***	4.45	(0.16)	28.17	<.001 ***	4.28	(0.12)	37.28	<.001 ***
Reference point other ideas	−0.52	(0.05)	−11.17	<.001 ***	−0.46	(0.05)	−8.92	<.001 ***	−0.28	(0.03)	−8.19	<.001 ***
Ownership bias	−0.49	(0.20)	-2.40	0.016 *	0.34	(0.21)	1.64	.102	0.40	(0.20)	1.99	.046 *
Task exposure	−0.06	(0.03)	−2.10	.036 *	−0.07	(0.03)	−2.45	.014 *	0.07	(0.03)	2.44	.015 *
Mean value experts	0.26	(0.04)	7.08	<.001 ***	0.39	(0.02)	18.26	<.001 ***	0.60	(0.02)	28.51	<.001 ***
Interaction												
*Task exposure * experts*	0.03	(0.06)	0.54	.592	0.05	(0.03)	1.52	.130	0.08	(0.03)	2.53	.011 *
Number of unique observations	1849				1858				1859			
Numner of observations	7396				7432				7433			
R-squared	0.03				0.09				0.19			

Note: Standard errors in parentheses. Significance levels indicated as follows: * $p < .05$, ** $p < .01$, *** $p < .001$.

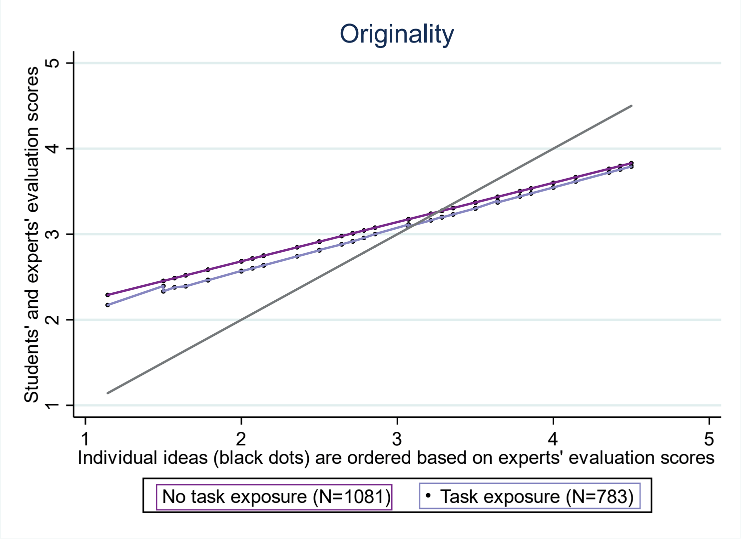
4

Figure 4.2.
Students' and experts' evaluation scores for each individual idea for (i) creativity, (ii) originality and (iii) feasibility

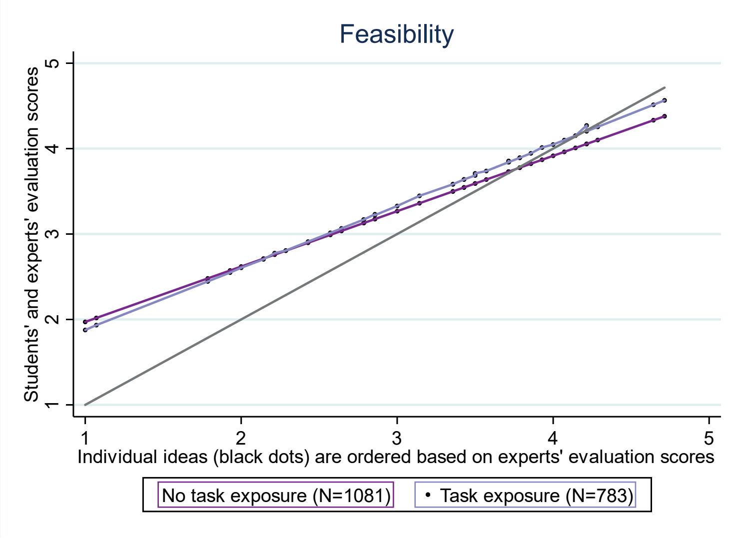
i.



ii.



iii.



4.4 Discussion

The main aim of the present study was to investigate the effect of task exposure on interpersonal idea evaluation. The results indicated that task exposure improves people's ability to recognize creative, original, and highly feasible ideas. To the best of our knowledge, this is the first study demonstrating the effect of task exposure on creative idea evaluation of others.

The results of this study showed that task exposure improved people's ability to accurately recognize creative and original ideas, and their ability to discriminate between highly feasible and less feasible ideas. Contrary to our expectations, task exposure did not improve people's ability to recognize (average) feasible ideas. This finding may be explained by the self-enhancement theory (Buehler, Griffin, & Ross, 1994; Dailey & Mumford, 2006; Jackson & Hogg, 2010). This theory refers to a general tendency of people to be overoptimistic about ideas when they are attached to a subject. As a result of this exaggerated optimism, people may come to think that solutions are more feasible to implement than they actually are. For example, Langholtz, Gettys, and Foote (1995) found that people tend to underestimate resource requirements when they have a certain amount of attachment with the product in question. In line with self-enhancement theory, we found that students with task exposure became more attached to the task. Due to this attachment, students might have become more sensitive to self-enhancement effects, causing them to overestimate the ease with which solutions can be implemented.

Next to the beneficial effects of task exposure on idea evaluation, we found other interesting findings worthwhile to discuss. First, we found a consistent pattern for the ratings of originality and creativity for both experts and students. In line with prior research, this can be explained by the fact that people consider the novelty criterion more important to determine the creativity of an idea than the feasibility criterion (Diedrich, Benedek, Jauk, & Neubauer, 2015; van Broekhoven, Cropley, & Seegers, 2020). As expected, our data shows a strong positive correlation between originality and creativity.[19] Thus, while the two evaluation criteria referred to different theoretical constructs, students perceived them as similar constructs. Even though the difference between both constructs was explicitly explained to students before the evaluation of ideas, this finding demonstrates that explanation in itself is not enough to teach people the difference between the two.

Second, while both content experts and creativity experts agreed on the assessment of ideas in terms of their originality and creativity, they differed in their assessment of feasibility. Content experts rated ideas lower on feasibility while creativity experts rated ideas higher on this aspect. As the main difference between content and creativity experts is their level of domain knowledge, this finding may be explained by the underpinnings of cognitive fixation reasoning theory (Dane, 2010; Lewandowsky, Little, & Kalish, 2007; Wiley, 1998). This theory asserts that people with high expertise may be more likely to focus on conventional expertise-related content, and are therefore more critical about feasibility aspects of products and ideas that are related to their domain of expertise than lay men. Moreover, Cseh and Jeffries (2019) described that consensual assessment technique (CAT) judges are required to be "knowledgeable" or "experts" in the domain to which the task is associated (p. 160). As such, creativity experts are not always suitable judges in the CAT, and, therefore, future studies should only consider experts with a certain degree of domain knowledge of the topic.

Several limitations should be noted. First, it can be argued that the beneficial effects of task exposure on idea evaluation were not due to task exposure, but to ownership bias. In this bias, people who have previously generated ideas and have to rate that same idea will favor *their* idea. However, we showed that there was (almost) no overlap between the generated ideas of the students and the ideas they

19 $r = 0.571$ ($p < .001$).

had to rate (4.5% of the students had to evaluate one or more ideas that they had earlier generated themselves, while 95.5% of the students never had to evaluate an idea that they had previously generated themselves). Moreover, we controlled for a potential 'ownership effect' by including a control variable in our analyses. As such, we attempted to convince the reader that the more accurate evaluations were due to more task exposure or familiarity, and not due to ownership bias.

Second, Myszkowski and Storme (2019) have recently proposed the Judge Response Theory (JRT) as an alternative framework for judge analysis instead of the Classical Test Theory (CTT). In JRT latent attributes – trait(s) and/or class(es) – of a product and of a judge are used as predictors of observed judgments. As such, while CTT reduces judges characteristics to random error and tries to eliminate it through averaging/summing, JRT includes judges characteristics in the measurement model (see Myszkowski & Storme, 2019 for more information). As such, it may be interesting for future research to include judges' characteristics in their measurement model.

Third, the present study compared two groups of students: (a) a group that had to generate ideas for a task for which they later also had to evaluate ideas (i.e., task exposure group); and (b) a group who had to generate ideas for a different task than they later had to evaluate ideas for (i.e., no task exposure group). Thus, both groups generated ideas, but for different kind of tasks. This was done to ensure that possible effects of task exposure could be causally investigated and to make sure that different evaluation outcomes could not be due to the pure effect of being exposed to a divergent thinking task before having to evaluate ideas alone (e.g. Berg, 2016). However, future research should add an extra group of students who have not been exposed to any idea generation task.

Furthermore, while the present research showed that task exposure improves people's ability to recognize the originality and creativity of ideas, and their ability to separate highly feasible ideas from highly unfeasible ideas, it is important to mention that the effects of task exposure were small in size (Cohen's d between .09 and .14). This is not surprising, given that this was only a minimal intervention that could be processed very quickly by people (four minutes on average). However, even though these effect sizes were relatively small, they are still big enough to make the difference in the daily practice of professionals who have to evaluate the creativity of other people's ideas (e.g. office manager, teacher). To illustrate in a more intuitive form what these effect sizes mean, Cohen's (1988) arcsine transformation convert them to percentage differences in creativity ratings. For example, imagine a manager in an innovative company who has to rate four ideas of his R&D department. Imagine that the selection of original ideas is done randomly: 50% chance to reject or select an original idea. As such, the approximate rate of originality rating in the no task exposure group is set at 50%. The mean effect size of .14 for task exposure is then equivalent to about a 7 percentage point differential, that is, a 57% for managers with task exposure to give ideas one originality score lower than managers without task exposure.[20] In this case, task exposure prevents the manager of accepting unoriginal ideas for further implementation. As such, the mean effect size of .14 could make all the difference in a creative idea being abandoned or further developed for implementation.

In summary, the present research shows that when exposed to the task, people's ability to recognize creative and original ideas, and their ability to discriminate between highly feasible and unfeasible ideas improves. These findings have implications for many areas, including advertising, marketing, and education. In these areas, professionals who are responsible for evaluating others' ideas are often purposely not exposed to the task at hand (e.g. creative directors, managers, teachers, and parents). However, the results suggest that professionals' ability to recognize original, creative, and highly feasible ideas improves when they are exposed to the task. In practice, this means that professionals would be provided with the opportunity to first individually generate ideas before evaluating others' ideas. By exposing professionals to the task at hand they may detect innovative ideas that could potentially change the world as we know it.

20 7% is deducted from the b-coefficient for the effect of task exposure on creativity rating: -0.07 * 100 (see Table 2).

5

To make or not to make? The role of expected implementation and personality in children's idea selection

Abstract

Worldwide, pedagogies that give children opportunities to develop creative and innovative competences have risen dramatically across primary schools. One of the core elements of these pedagogies is that children are asked to work on transforming their ideas into tangible and physical products. In this study, we examined whether expected implementation of ideas into tangible products affects the selection of original and feasible ideas. Furthermore we investigated whether personality traits moderates this relationship. To this end, 403 grade 6 students (aged 10-13) from 13 primary schools in the Netherlands were asked to select two innovative ideas to improve the use of a stuffed toy elephant with or without the expectation to actual transform these ideas into crafts. The results showed that children who expected implementation were less likely to select original ideas, but more likely to select feasible ideas than children who had no such expectation. Moreover, children in the implementation condition focused more on feasibility than originality when selecting innovative ideas. Regarding the moderating

role of personality, we found that children scoring high on conscientiousness were more inclined to select feasible ideas, even though they were instructed to select innovative ideas and did not expect idea implementation. Together, the results highlight the importance for educators to carefully consider whether or not practical components should be part of assignments, and to tailor instruction in assignments to the individual needs of children.

This chapter is joint work with Barbara Belfi and Lex Borghans. The research was supported by the Educatieve Agenda Limburg and the Kindante board of primary schools. We thank Maurice Thelen and Sylvie Beckers for their support in the experimental design, and their expertise on how to integrate this in the grade 6 curriculum. We thank the participants of the 2020 MIC Conference for their valuable comments and suggestions on earlier versions of this paper.

5.1 Introduction

Today, primary schools spend more and more attention to the development of creative and innovative thinking competences (van Broekhoven, Cropley, & Seegers, 2020; OECD, 2019). One of the reasons for this growing interest is that a large share of the tasks carried out by people in their daily jobs may be automated by robots or machines in the foreseeable future. It is expected that the tasks that will be left over for human labor will be those that cannot be easily transformed in coding language and are thus in need of human higher-order competences, such as creativity and innovation (Allen, Belfi, & Borghans, 2020; International Labour Organization, 2017; World Economic Forum, 2016; 2018). To this end, constructivist pedagogies that aim to enhance children's' creativity and innovation competencies have risen dramatically across primary schools (Jeffrey & Craft, 2004; Said-Metwaly, Fernández-Castilla, Kyndt, Van den Noortgate, & Barbot, 2020).

In the fields of psychology and education, *creativity* is often described as the generation of ideas that are both original and useful whereas *innovation* is described as the selection and actual implementation of these original and feasible ideas (Anderson, De Dreu, & Nijstad, 2004; Barbot, Besançon, & Lubart, 2015; Runco & Jaeger, 2012). In this definition, originality refers to the novelty of an idea and usefulness to its feasibility and valuableness. Particularly in constructivist learning methods, much attention is paid to enhancing students' creative and innovative competencies (McLoughlin & Lee, 2008; Piagèt, 1964; Vygotsky, 1980). These learning methods come in many shapes and forms (e.g. Montessori education, design thinking, problem-based or research-based learning), but all share a common focus on stimulating students' imagination and curiosity by means of self-discovery (e.g. Besançon & Lubart, 2008; Heise, Böhme, &, Körner, 2010; McLoughlin & Lee, 2008; Piagèt, 1964; Vygotsky, 1980).

An important characteristic of constructive teaching methods is, furthermore, that children are frequently asked to work on transforming their ideas into tangible and physical products (Beghetto & Kaufman, 2010; Cardarello, 2014; Davies et al., 2013). These products can be presented in different forms such as essays, drawings or craft projects. In primary schools, particularly drawing and crafting are popular means to asses creativity and innovation competencies as they can be seen as the universal language of all children, irrespective of their age, nationality or intelligence (Alfonso-Benlliure & Santos, 2016; Amabile, 1996; Driessnack, 2005; Jellen, Urban, & Quarterly, 1986). Nonetheless, there are hints in the literature that the expected transformation of original ideas into tangible products might impact final product innovation in a negative way, because highly original ideas are often perceived as infeasible to implement (e.g. Baer, 2012; Mueller, Melwani, & Goncalo, 2012; Rietzschel, Nijstad, & Stroebe, 2010; Shalley & Zhou, 2008; West et al., 1990). Therefore, children may be prone to select more common ideas when they have to transform them into tangible products. Hence, the expectation of having to implement an idea into a product may result in the selection of more common ideas to be transformed into products, while the original ideas are abandoned for further implementation. The aim of this study is therefore to investigate whether expected implementation of ideas affects children's selection of original and feasible ideas.

The selection of original and feasible ideas for transformation in tangible products may further be moderated by children's personality (Toh & Miller, 2016b). For example, children scoring high on the personality trait *openness to experience* may be more eager to select original ideas that are unconventional or potentially risky than children scoring low on this personality trait. In contrast, children who are highly *conscientiousness* may find naturally more comfort in mundane ideas rather than original ideas. As such, children with different personality traits might select different kind of ideas when they expect to implement their ideas. Therefore, the second aim is to investigate whether children's personality moderates the relationship between expected idea implementation and the selection of original and feasible ideas.

5.2 Effect of expected implementation on idea selection

Little is known about the role of expected implementation of ideas on children's selection of original and feasible ideas. One of the reasons why expected implementation may have an impact on children's idea selection, is because they may be afraid that it will be too difficult to transform their idea in a feasible product or that they may be ridiculed by others seeing their innovative product. Yuan and Zhou (2008) investigated the extent to which expected evaluation by others affected the idea selection of undergraduate students. In their study, seventy-three students were asked to generate and select ideas, and modify them into final solutions. During idea selection, students were reminded that their final solutions should be creative (i.e., both original and feasible). Yuan and Zhou found that students who expected evaluation modified their ideas more by making them feasible to implement. Alternatively, this did not occur among students who did not have such an expectation. Moreover, students who expected evaluation of their ideas selected more creative ideas than students who did not expect evaluation.

The main explanation provided for these results is that extrinsic constraints, such as expected evaluation, may have a facilitating effect on the performance of algorithmic tasks (Amabile & Pratt, 2016; McGraw, 1978). Algorithmic tasks are tasks where the paths to the solution and methods to reach that solution are well-mapped and straightforward. For example, McGraw (1978) illustrated that while it might not be easy to find a solution for a complex multiplication task, the solution to this task can be found by means of a well-defined procedure. Idea selection can also be considered as an algorithmic task that has a well-defined procedure because children know the standards to evaluate their ideas with and know what the rules are (i.e., materials and criteria of ideas) to come to a final solution. These rules provide a roadmap, or checklist for idea selection. For such algorithmic problems, extrinsic constraints - such as expected evaluation or implementation - may actually serve as a reinforcement to increase effectiveness. A more effective idea selection process means that children will spend more attention to the potential feasibility of ideas. As such, we hypothesized that *children who expect implementation will select more feasible ideas than children who do not expect implementation* (H1a).

However, as expressed before, a final solution will not be considered creative or innovative unless it is both original and feasible (Amabile, 1996; Runco & Jaeger, 2012 for reviews). The originality and usefulness trade-off describes that these two components cause tension in idea selection, because original ideas are often perceived as useless or infeasible (Manske & Davis, 1968; Mueller et al., 2012; Rietzschel, Nijstad, & Stroebe, 2006; Rietzschel et al., 2010).[21] The expectation of idea implementation may reinforce children's persistent efforts and willingness to comply with necessary rules and standards. However, original ideas often do not comply with the fixed rules and standards, and, therefore, children are less likely to focus on idea originality when they expect idea implementation. Drawing on previous research findings, we hypothesized that *children who expect implementation will select less original ideas than children who do not expect implementation* (H1b).

As such, we predicted that expected implementation would have contradictory effects on the two dimensions of final product innovation. The expected implementation ensures that children will consider both the originality and the feasibility aspects in their idea selection. Moreover, they are encouraged to put more emphasis on feasibility while children without this extrinsic constraint may be more likely to turn a blind eye to feasibility. Yet, a merely original idea is not innovative, because a final product will not be considered as innovative unless it is both original

21 In line with the originality and usefulness trade-off, our data shows a statistical significant negative correlation between originality and feasibility (r = -0.357).

and feasible. As such, we hypothesized that *for children who expect implementation, compared to children who do not expect implementation, feasibility will increase more than originality will decrease* (H2).

5.3 The moderating role of personality

One of the best-established personality models is the five-factor model (or 'the Big Five'; Costa & McCrea, 1992; Digman, 1990; John & Srivastava, 1999), which encompasses five dimensions: openness to experience, conscientiousness, agreeableness, extraversion, and emotional stability (or neuroticism). Generally, personality traits have been strongly linked to creativity and innovation (e.g. Batey & Furnham, 2006; Feist, 1998; Woods, Mustafa, Anderson, & Sayer, 2018). Specifically, openness to experience, conscientiousness and agreeableness have been linked to idea selection (Reilly, Lynn, & Aronson, 2002; Toh & Miller, 2016a).

Openness to experience represents the extent to which people are curious, imaginative, broad-minded, and non-traditional (Mount & Barrick, 1995). As such, children who are open to experience are more flexible in embracing original ideas even though these ideas may be untested (Goldberg, 1990). In contrast, children with lower levels of openness to experience tend to find more comfort in familiar ideas rather than original and new ones (Choi, 2004; George & Zhou, 2001). Therefore, we hypothesized that *the effect of expected implementation on choosing feasible rather than original ideas to craft, will be stronger for children with low levels of openness to experience than children with high levels of openness to experience* (H3a).

Furthermore, while people with lower levels of conscientiousness tend to be careless, quit easily, distractible and unreliable (Goldberg, 1990), people with higher levels of conscientious tend to be attentive, persistent, orderly and neat (Shiner & Caspi, 2003). As such, conscientious children have a focus on "doing things right" and, therefore, we expect that they would respond better to the instructions given to them. Hence, we hypothesized that *the effect of expected implementation on choosing feasible rather than original ideas to craft, will be stronger for children with high levels of conscientiousness than children with low levels of conscientiousness* (H3b).

Finally, children demonstrating lower levels of agreeableness tend to be more selfish, manipulative, and rude (Goldberg, 1990). In contrast, children scoring high on agreeableness tend to care about others' feelings and avoid conflict with other children (Shiner & Caspi, 2003). Yet, original ideas are often regarded as challenging the status quo and thus disrupting interpersonal relations and processes endorsed by others, which can cause conflict with others (Choi, 2007; Lim & Choi, 2009). As such, we expect that children with high scores on agreeableness may be less inclined to select original ideas for implementation, as these new and unexpected ideas may be in conflict with the ideas of other children. Therefore, we expect that *the effect of expected implementation on choosing feasible rather than original ideas to craft, will be stronger for children with high levels of agreeableness than children with low levels of agreeableness* (H3c).

5.4 Methods

5.4.1 Sample

Before data collection, a priori power analysis revealed that a minimum of 388 participants is required to obtain a statistical power of 0.99 with a (independent) t-test (G*Power 3.1; Faul, Erdfelder, Buchner, & Lang, 2009). We planned to recruit slightly more participants to compensate for drop outs due to potential technical issues (e.g. problems with internet connection in the school).

The data were collected from 403 children from thirteen primary schools in Limburg, the most southern province of the Netherlands, between February and June 2019. The children (49.9% girls, 50.1% boys) attended Grade 6 (last grade of primary school) and were 11.6 years of age on average (SD = 0.48). [22] From 353 children (88% of the total sample), personality traits were measured. This subsample of 353 children will be used to investigate whether the personality traits openness to experience, conscientiousness and agreeableness moderate the relationship between expected implementation and idea selection.

The current study was approved by the Ethical Review Committee Inner City Faculties from Maastricht University (approval code: ERIC_090_14_06_2018). Parental consent was obtained through passive consent wherein parents were informed and asked to contact the research team if they did not approve their child to participate. From two children, the research team received withdrawal from participation.

22 The effect of expected implementation on idea selection was tested among 10-13 aged children, because recent research found children around this age begin a trend of increasing conformist thinking that continues through high school (Kim, 2011; Said-Metwaly, Fernández-Castilla, Kyndt, Van den Noortgate, & Barbot, 2020; Jastrzebske & Limont, 2017).

5.4.2 Procedure and experimental design

The experiment took place at the children's own school. In the school's computer lab, children filled out an online assignment (see Appendix A). The assignment consisted of three tasks and took about 20 minutes to complete (see Figure 5.1). Both the experiment leader and the teacher, walked around to answer questions about possible ambiguities. In the online assignment children were first asked to evaluate ideas on their feasibility and originality for improving a stuffed toy elephant, such as enlarge the toy elephant or create a toy elephant that is able to spit fire. Next, all children were asked to generate as many ideas as possible for toys for monkeys in the zoo.[23] After this task, a randomized design was implemented with two experimental conditions: *expected implementation* and *non-expected implementation.* Children in the expected implementation condition (N=201) were told, through the written instruction on the computer screen, to expect future implementation of their selected ideas:

> "A toy factory needs your help! The toy factory makes toy animals, such as elephants, dogs, rabbits and so on. They would like to receive innovative ideas to change a toy elephant. They will first test these ideas on a toy elephant made of paper. **You will craft these ideas.**"

In contrast, children in the non-expected implementation condition (N=202) were told:

> "A toy factory needs your help! The toy factory makes toy animals, such as elephants, dogs, rabbits and so on. They would like to receive innovative ideas to change a toy elephant. They will first test these ideas on a toy elephant made of paper. **You will <u>NOT</u> craft these ideas, because you will be crafting ideas for monkey toys.**"

After reading these instructions, children were asked to select five innovative ideas to improve a toy elephant. [24] From those five ideas, they had to select the two most innovative ideas.[25] Thus, the manipulation was that some children were told that

23 This idea generation task was implemented to prevent memory effects in the idea selection task that followed.

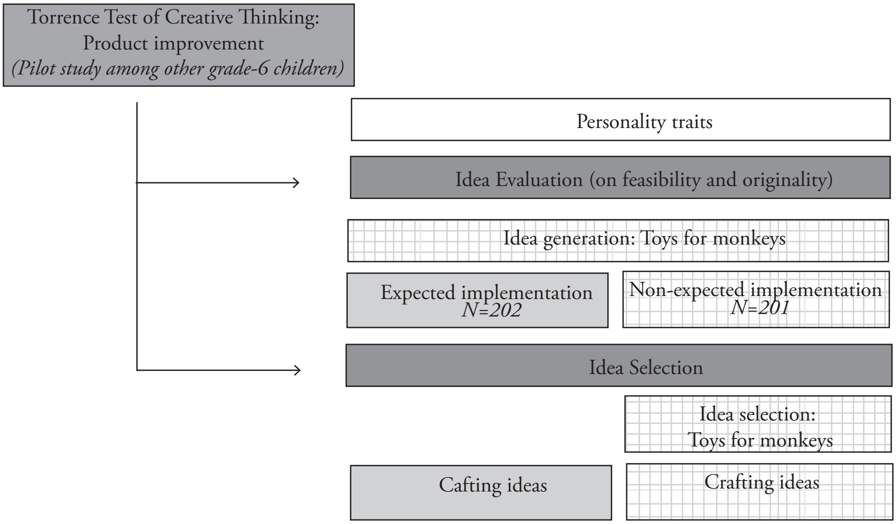
24 The pilot study revealed that grade-6 children had problems with understanding the meaning of 'creative'. In contrast, they did understand the meaning of 'innovative'. Since children's understanding of the term 'innovative' appeared to be synonymous to the meaning of 'creative', the former term was used in our communication to children rather than the latter.

25 The results are robust for five innovative ideas and for two innovative ideas.

5

they later had to implement (i.e., craft) these ideas, while other children were told that would not have to implement these ideas. The non-expected implementation condition received an additional question where they selected one idea for a toy for monkeys in the zoo to craft in the classroom. Immediately after the online assignment, children went back to their classroom where they had to craft the two toys from their final selected ideas of the third task. After this crafting exercise, the experiment ended and the children went back to their normal school program.

Figure 5.1
Experimental design



5.4.3 Measures

Idea selection task. As part of this study, 36 grade-6 pupils from a previous cohort had generated ideas to improve a stuffed toy elephant as part of the Torrance Test of Creative Thinking (TTCT; see Torrance, 1974). This product improvement task resulted in 438 ideas. This number was reduced to 369 by excluding ideas that involved non-play uses, such as make the elephant alive or use it as a pincushion. The ideas were then further reduced to a list of 63 ideas by excluding ideas that were similar (i.e., the ideas 'make it bigger' and 'making a XL elephant' were collapsed into one idea, 'enlarge the toy elephant'). Next, the remaining 63 ideas were rated by seven experts (i.e., four primary school teachers and three creativity researchers).

The experts were instructed to rate each idea on feasibility and originality using a Likert scale ranging from 1 (not at all feasible/original) to 5 (very feasible/original).[26] To reduce the list of 63 ideas even further, a set of twenty ideas was selected to be presented to the children in the experiment. The interrater reliability was high: The overall intraclass correlation coefficient (ICC, two-way random, consistency analysis) was .90, and the single interrater reliabilities were excellent (feasibility ICC = .94 and originality ICC = .86). By averaging the scores of the seven experts, each single idea received a feasibility and originality score.

Idea selection performance. This performance was measured by two indicators: feasibility and originality. The first indicator of idea selection performance, feasibility, was measured by children's own-rated feasibility of the two selected ideas. Prior to idea selection, children were asked to evaluate twenty ideas for the stuffed toy elephant on their feasibility. Similarly as in Charles and Runco (2001), children were asked to select one of five faces (frown = *very difficult to craft this idea*; slight frown = *difficult to craft this idea*; no expression = *not difficult but also not easy to craft this idea*; slight smile = *easy to craft this idea*; smile = *very easy to craft this idea*) that best showed what they thought of the idea.[27]

The second indicator of idea selection performance, originality, was measured by children's own-rated originality of the two selected ideas. The originality of an idea was operationalized as the degree to which children thought that other children would have a similar idea on a scale from 1 to 10. Thus, their task was to estimate a hypothetical number of children able to generate a given idea (Charles & Runco, 2001).[28] Next, we transformed this scale to a 5-point Likert scale to compare this performance measures with that of children's feasibility ratings (range from 1 to 5). This was done by first dividing the rating by two (range from .5 to 5). Next, we

26 Additionally, the experts were asked to estimate how many children, out of 10, they thought could have generated each idea (i.e., statistical infrequency scale). This measure is often used in creativity research to determine the originality of ideas (e.g. Charles & Runco, 2001; Grohman, Wodniecka, & Klusak, 2006; Runco & Dow, 2004; Silvia, 2008; Wagner, 1993; Wallach & Kogan, 1965). Moreover, the pilot study showed that children are well able to assess the originality of ideas in this way (Charles & Runco, 2001). The validity of this scale was supported by the finding that among experts, the scores on this 'statistical infrequency scale' and the 5-point Likert scale were highly correlated ($r = 0.841$).

27 As a robustness check, the average rating of the seven experts was used as well to measure idea selection performance for feasibility.

28 This is a much-used measure in creativity research to determine the originality of ideas, particularly in research with children (e.g. Charles & Runco, 2001; Grohman, Wodniecka, & Klusak, 2006; Runco & Dow, 2004; Silvia, 2008; Wagner, 1993; Wallach & Kogan, 1965).

reversed these values so that higher values indicated a higher originality (range from 1 to 5).[29]

Personality traits. Personality traits were measured using the 50-item version of the International Personality Item Pool (IPIP; Goldberg et al., 2006). For each personality trait, children received ten adjusted statements suitable for children (presented in a random order). Sample statements include: "I am bursting with ideas" and "I am always prepared". Children rated how well each statement describes themselves on a Likert scale, ranging from 1 (very inaccurate) to 5 (very accurate). Scale reliability (Cronbach's alpha) in this study was good: openness to experience ($\alpha = .75$), conscientiousness ($\alpha = .80$), agreeableness ($\alpha = .74$), extraversion ($\alpha = .71$) and emotional stability ($\alpha = .82$).

Manipulation check. To check whether the manipulation was effective, one item in the questionnaire was used to check whether the children experienced the pressure of the expected implementation. Half of the children in both the expected and non-expected implementation condition were asked whether they thought that they had to craft ideas to improve a toy elephant or a toy for monkeys in the zoo. The children responded with the toy elephant or a toy for monkeys. This manipulation check was done to determine whether children in the experimental condition actually realized that they had to implement the ideas they selected.

5.4.4 Statistical analysis

To test hypotheses 1 – about differences in feasibility and originality ratings – we performed univariate ordinary least squares regressions with children's average rating for the two selected ideas as outcome variables, separately for feasibility and originality:

$$Y_{ija} = \beta_0 + \beta_1 \text{ treatment}_{ij} + e_{ij}$$

where Y_{ija} is the outcome of child i in group j ($j = 0$ for control, $j = 1$ for treated) and a refers to feasibility or originality rating for the two selected ideas. As such, Y_{ija} indicates the children's average own rating of the two selected ideas, separately for feasibility and originality. The variable of interest, $treatment_{ij}$, is a binary variable that takes value 1 for children who expected implementation of their selected ideas, and zero for children who did not expect implementation of their selected ideas. Lastly,

29 As a robustness check, the average rating of the seven experts was used as well to measure idea selection performance for originality.

e_{ij} is a normally distributed residual with zero mean and constant variance σ_e^2 (e.g. Porter & Raudenbush, 1987; Reichardt, 1979; van Breukelen, 2013).

To test hypothesis 2 – about the trade-off between feasibility and originality - an empirical model should simultaneously estimate the effects of children's feasibility and originality ratings on their selection of ideas. This allows us to better understand children's considerations in relation to feasibility and originality when taking a decision about the creativity an idea. A statistical model that suits this type of decision-making process is the conditional logit model, proposed by McFadden (1974). In brief, each child chooses two ideas of a set of eighteen ideas to improve a stuffed toy elephant. The probability that a child *i* chooses *k* among *j* alternatives is:

$$\Pr(i\ chooses\ k) = \Pr(V_{ik} > V_{ij}) \quad \forall\ j \neq k, j = 1, \ldots, J$$

In general, the utility of an alternative *j* for the child *i* is given by:

$$V_{ij} = x_{ij}\beta + v_{ij} \qquad i = 1, \ldots, n; j = 1, \ldots, J$$

where x_{ij} the feasibility and originality ratings are varying across ideas. These ratings interact with treatment (i.e., expected implementation or non-expected implementation).[30] Specifically, the interaction terms included were:

- $Treatment_i * Feasibility_{ij}$
- $Treatment_i * Originality_{ij}$

Results are reported in odds ratios. These should be interpreted as the proportional change in the odds of child *i* selecting idea *k* for a unit increase in the treatment variable, holding all other variables constant. This means that we can draw conclusions about the probability for children in the expected implementation condition and children in the non-expected implementation condition to select an idea given its feasibility and originality. It is important to be clear about what is meant by 'change in the odds' in these models. It is based on the number of children making a particular choice (i.e., selection of two ideas), whilst accounting for the number of alternatives available within that choice set (i.e., 18 other ideas). Thus, when we say that the probability

30 All variables must vary across ideas (or alternatives) in order to achieve identification in the conditional logit model. Therefore, as treatment is alternative-invariant, it can only be included in the model as an interaction(s) with the characteristics of the alternatives (i.e., feasibility and originality).

of selecting a feasible idea is higher among children in the expected implementation condition than children with no such expectation, this is after accounting for the number of ideas in the total set.

To test hypotheses 3 – about moderation effects with personality traits – we performed multivariate ordinary least squares regressions with children's average rating for the two selected ideas as outcome variables, separately for feasibility and originality:

$$Y_{ija}=\beta_0+\beta_1\, treatment_{ij}+ \beta_2\, P_{ij}+ \beta_3\, treatment{*}P_{ij}+ e_{ij},$$

where Y_{ija} is the outcome of child i in group j (j = 0 for control, j = 1 for treated) and a refers to feasibility or originality. Additional to the univariate ordinary least squares regression, P_{ij} is an interval variable in which children can score 1 (not at all) to 5 (very) on a personality trait, separately for openness, conscientiousness, extraversion, agreeableness, and emotional stability. The variable of interest, $treatment{*}P_{ij}$, is an interaction term of the binary treatment variable with an interval personality trait variable.

5.5 Results

5.5.1 Descriptive Statistics and Randomization Check

Before analysing treatment effects, the descriptive statistics (see Table 5.1) of demographic background and personality traits were calculated. As a randomization check, the differences in demographic background variables (i.e., gender, age, migration background and parental socioeconomic status) and personality traits (i.e., openness to experience, conscientiousness, extraversion, agreeableness and emotional stability) between the two experimental conditions were tested for statistical significance. No statistically significant differences between the conditions emerged, based on two-tailed t-test tests with an a level of 5%: t (gender) = 0.08, p = -1.75; t (age) = 0.86, p = -0.18; t (migration background) = 0.30, p = 0.77; t (parental low SES) = -0.47, p = 0.64; t (parental middle SES) = 1.42, p = 0.16; t (parental high SES) = -1.10, p = 0.27; t (openness to experience) = -1.66, p = 0.10, t (conscientiousness) = 0.17, p = 0.87, t (extraversion) = -0.73, p = 0.46, t (agreeableness) = -0.33 , p = 0.74, and t (emotional stability) = 0.55, p = 0.58 .

Table 5.1
Descriptive statistics conditional on experimental group

	Expected implementation					Non-expected implementation						
	N	*Mean*	*SD*	*Min*	*Max*	*N*	*Mean*	*SD*	*Min*	*Max*	*t-value*	*p-value*
Demographic information												
Gender (ref.=boys)	201	0.54				202	0.46				0.08	-1.75
Age (in years)	201	11.59	0.49	10.65	13.02	202	11.58	0.48	10.37	13.16	0.86	-0.18
Migration background (ref.=Dutch)	195	0.03				196	0.03				0.30	0.77
Parental socioeconomic status (ISCED)												
Low	122	0.08				121	0.07				-0.47	0.64
Middle	122	0.28				121	0.36				1.42	0.16
High	122	0.64				121	0.57				-1.10	0.27
Personality traits												
Openness to experience	180	3.51	0.60	1.50	5.00	173	3.39	0.66	1.25	5.00	-1.66	0.10
Conscientiousness	180	3.51	0.68	1.88	4.89	173	3.52	0.67	1.78	5.00	0.17	0.87
Extraversion	180	3.25	0.62	1.44	4.67	173	3.21	0.62	1.33	4.67	-0.73	0.46
Agreeableness	180	3.94	0.58	2.14	5.00	173	3.92	0.58	2.43	5.00	-0.33	0.74
Emotional stability	180	2.90	0.74	1.11	4.44	173	2.94	0.71	1.00	4.78	0.55	0.58
Manipulation check	115	0.93	0.26			98	0.10	0.30			-21.60	0.00 ***

Note: The table reports the number of students, mean values, standard deviations, minimum and maximum value for demographic information and personality traits. There were no differences in the number of students per school. Significance levels are indicated as follows: * $p < .05$. ** $p < .01$. *** $p < .001$.

5.5.2 Manipulation Check

To check whether children in the expected implementation condition indeed expected that they had to craft their selected ideas, and children in the non-expected implementation condition did not have such expectation, we conducted a univariate one-way ANOVA. ANOVA results showed that the manipulation was successful ($F(1, 211) = 466.49$, $p < .001$, $\eta^2 = 0.69$). As shown in Table 5.1, children in the expected implementation condition ($M = 0.93$, $SD = 0.26$) had significantly higher levels of expectation to craft ideas for a toy elephant than participants in the non-expected implementation condition ($M = 0.10$, $SD = 0.30$).

5.5.3 Hypothesis Testing

Hypothesis 1a predicted that children who expected implementation would select more feasible ideas, while hypothesis 1b predicted that children who expected implementation would select less original ideas than children who did not expect implementation. Univariate ordinary least squares regressions were carried out to test if expected implementation significantly predicted the average feasibility and originality of the two selected ideas.

As shown in Table 5.2, expected implementation had a significant effect on the feasibility of the two select ideas ($b = 1.16$, $p < .001$, $\eta^2 = 0.21$). As predicted, children who expected implementation of ideas selected more feasible ideas ($M = 3.28$, $SD = 1.14$) than children who did not have such expectations ($M = 2.12$, $SD = 1.12$; see Figure 5.2). Hence, we found support for hypothesis 1a.[31]

In addition, Table 2 shows that expected implementation significantly predicted children's average originality of the selected ideas ($b = -0.43$, $p < .001$, $\eta^2 = 0.04$). As expected, children who were free from expected implementation selected significantly more original ideas ($M = 3.79$, $SD = 1.18$) than children who expected implementation of ideas ($M = 3.36$, $SD = 1.12$; see Figure 5.2). As such, we found support for hypothesis 1b as well. [32]

31 In line with these findings, the average rating of the seven experts as outcome variables showed similar results (see Table B.1 and Figure B.1 in Appendix B).

32 In line with these findings, the average rating of the seven experts as outcome variables showed similar results (see Table B.1 and Figure B.1 in Appendix B).

Table 5.2.
Ordinary Least Squares (OLS) regression with children's average rating for the two selected ideas, separately for feasibility and originality

	Feasibility				Originality			
	Model 1a b-coefficient	η^2	Model 2 b-coefficient	η^2	Model 1b b-coefficient	η^2	Model 2 b-coefficient	η^2
Constant	0.96 ***		1.32		4.23 ***		4.89 ***	
	(0.18)		(0.75)		(0.18)		(0.79)	
Main effect								
Expected implementation	1.16 ***	0.21	2.33 *	0.01	-0.43 ***	0.04	-1.64	0.01
	(0.11)		(1.08)		(0.12)		(1.14)	
Personality traits								
Openness to experience			-0.08				0.09	
			(0.14)				(0.15)	
Conscientiousness			0.37 **				-0.31 *	
			(0.14)				(0.15)	
Extraversion			0.03				-0.03	
			(0.14)				(0.15)	
Agreeableness			-0.04				-0.04	
			(0.17)				(0.18)	
Emotional stability			-0.10				-0.01	
			(0.12)				(0.13)	
Interaction with personality								
*Implementation*openness*			-0.05				0.17	
			(0.20)				(0.22)	
*Implementation*conscientiousness*			-0.54 **				0.38	
			(0.19)				(0.20)	
*Implementation*extraversion*			0.17				0.08	
			(0.20)				(0.22)	
*Implementation*agreeableness*			0.04				-0.28	
			(0.23)				(0.24)	
*Implementation*emotional stability*			0.11				0.02	
			(0.17)				(0.18)	
R-squared	0.21		0.27		0.04		0.07	
F-statistics	106.37		11.53		14.31		2.4	
Number of observations	403		353		403		353	

Note: Standard errors in parentheses. Effect sizes were calculated as partial eta squares (η^2) where 0.01 is considered to be a small effect size, 0.09 medium effect size and 0.25 a large effect size. Significance levels indicated as follows: * $p < .05$, ** $p < .01$, *** $p < .001$.

5

Figure 5.2
Mean feasibility and originality of two selected ideas (children's rating)

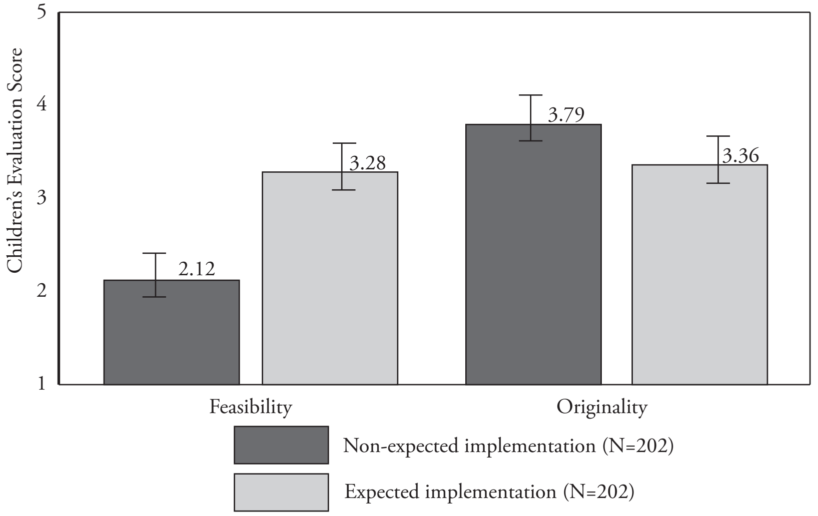


Figure 5.1 shows that the feasibility level of the final two selected ideas increased more than originality decreased for children who expected implementation, compared to children who did not expect implementation. To test hypothesis 2 – about the trade-off between feasibility and originality - conditional logit models were conducted to simultaneously estimate the effects of children's feasibility and originality rating on their selection of ideas. This allows us to better understand what children's considerations are in relation to feasibility and originality when taking a final decision about the creativity of an idea.

As shown in Table 5.3, the conditional logit confirmed the finding that expected implementation positively affected the likelihood of a child to select feasible ideas, while it negatively affected the likelihood of a child to select original ideas. More specifically, the expected implementation of ideas increased the probability of choosing a feasible idea with 103%, while the probability of choosing an original idea reduced with 14%.[33] Hence, the conditional logit model showed that the feasibility level of the two final selected ideas increased more (with 103%) than originality decreased (with 14%) for children who expected implementation, compared to children who do not expect implementation. Hence, we found support for hypothesis 2.

33 This was calculated based on Model 3 in Table 5.3 (100 – 86 = 14%).

Table 5.3

Idea selection conditional on alternative ideas (odds ratios and Z-statistics)

	Model 1		Model 2		Model 3	
Feasibility	0.60	***			0.62	***
	(-11.05)				(-10.33)	
Treatment*Feasibility	2.07	***			2.03	***
	(12.29)				(11.78)	
Originality			1.30	***	1.20	***
			(5.94)		(4.13)	
Treatment*Originality			0.77	***	0.86	*
			(-4.41)		(-2.52)	

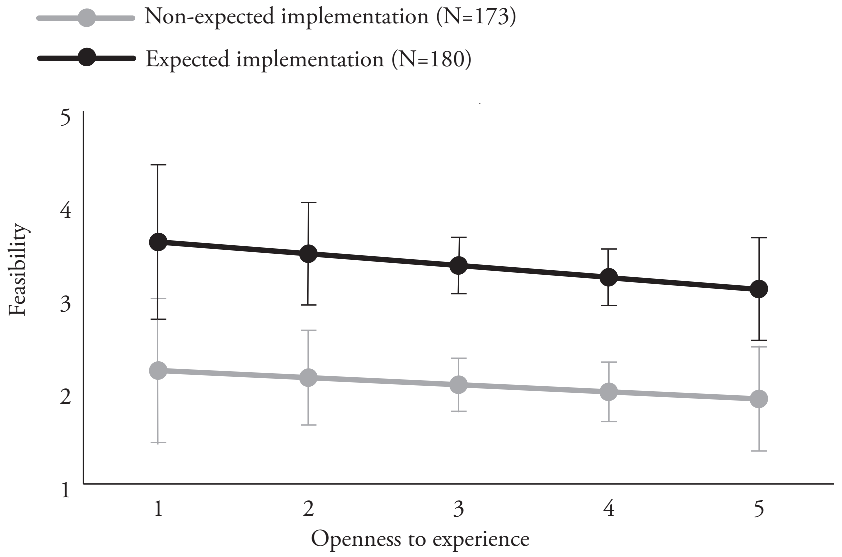
Note: Z-statistics are reported in the parentheses to indicate statistical significance. Effects are interpreted as the probability favoring idea *k* is multiplied with a one-unit increase in that variable. Estimates greater than one are considered positive effects, while estimates smaller than one are considered negative effects. *** Statistical significance at the 0.10 percent level (Z-statistics > 3.10), ** Statistical significance at the 1 percent level (Z-statistics > 2.58), * Statistical significance at the 5 percent level (Z-statistics > 1.96).

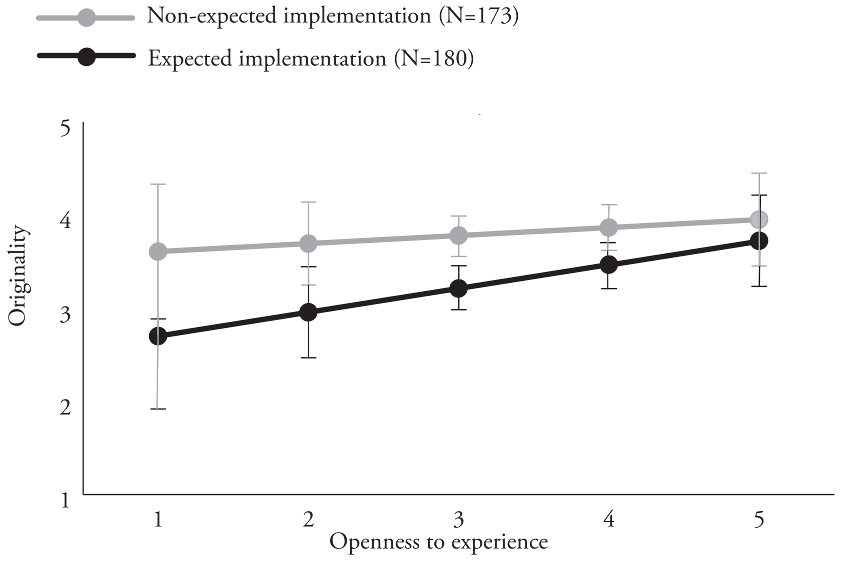
To fully understand the relationship between expected implementation and idea selection, we also examined whether the personality traits openness to experience, conscientiousness or agreeableness moderated this relationship. Hypotheses 3a predicted that the effect of expected implementation on choosing feasible rather than original ideas to craft, would be stronger for children with low levels of openness to experience than children with high levels of openness to experience. As shown in Table 5.2, multivariate ordinary least squares regression suggest that the main effect of expected implementation on feasibility is significant (b = 2.33, p < .05, η^2 = 0.01), but the interaction effect with openness to experience is not significant (b = -0.05, p > .10; see Table 5.2). Furthermore, neither the main effect of expected implementation on originality is significant (b = -1.64, p > .10, η^2 = 0.01), nor the interaction effect with openness to experience (b = 0.17, p > .10; see Table 5.2). To illustrate the nature of this interaction, Figure 5.3a shows that the effect of expected implementation on idea selection is not different for children high or low on openness to experience. As such, hypothesis 3a is rejected.

5

Figure 5.3a

Interaction between expected implementation and openness to experience for (i) feasibility and (ii) originality

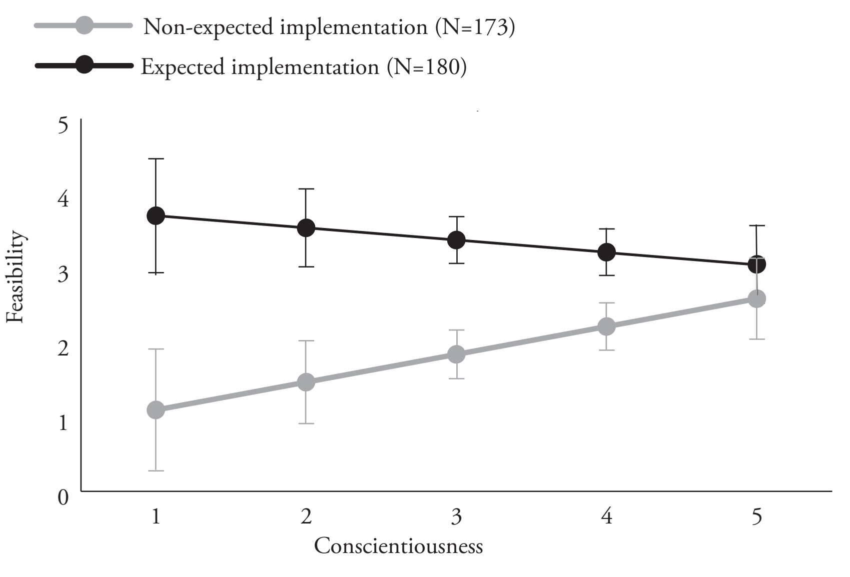


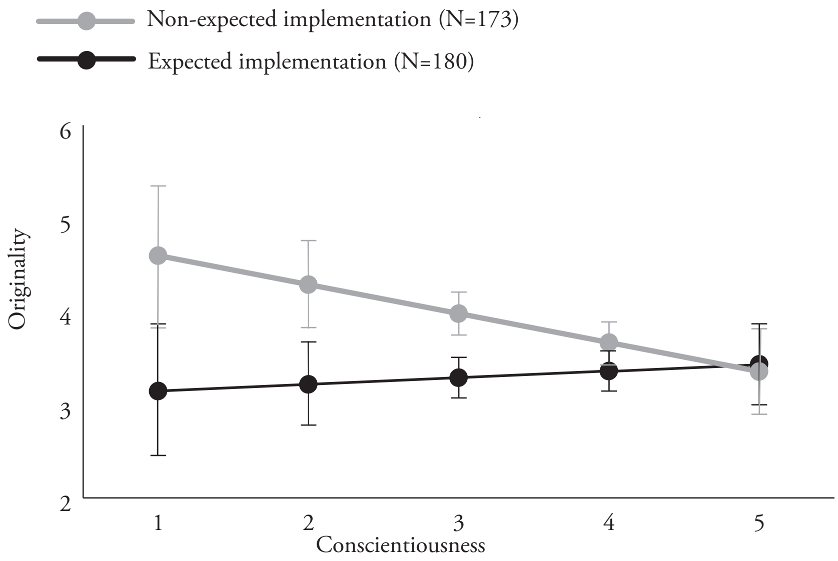


Hypotheses 3b predicted that the effect of expected implementation on choosing feasible rather than original ideas to craft, would be stronger for children with high levels of conscientiousness than children with low levels of conscientiousness. Multivariate ordinary least squares regression suggested that the main effect of expected implementation on feasibility is significant (b = 2.33, p < .05, η^2 = 0.01), and that the interaction effect with conscientiousness is significant as well (b = -0.54, p < .01; see Table 5.2). However, neither the main effect of expected implementation on originality is significant (b = -1.64, p > .10, η^2 = 0.01), nor the interaction effect with conscientiousness (b = 0.38, p > .10; see Table 5.2). To illustrate the nature of this interaction, Figure 5.3b shows that the effect of expected implementation on idea selection in terms of feasibility is different for children high or low on conscientiousness. Hence, we found partial support for hypothesis 3b.

Figure 5.3b

Interaction between expected implementation and conscientiousness for (i) feasibility and (ii) originality

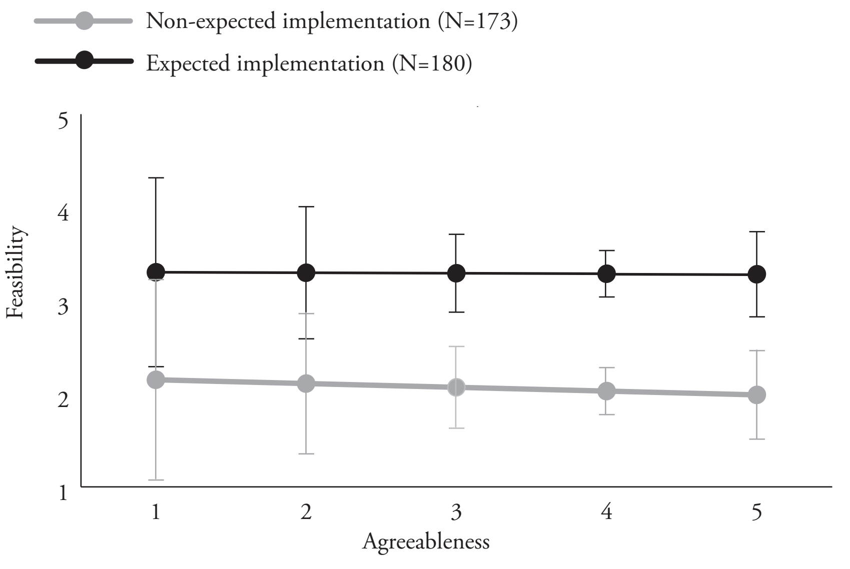


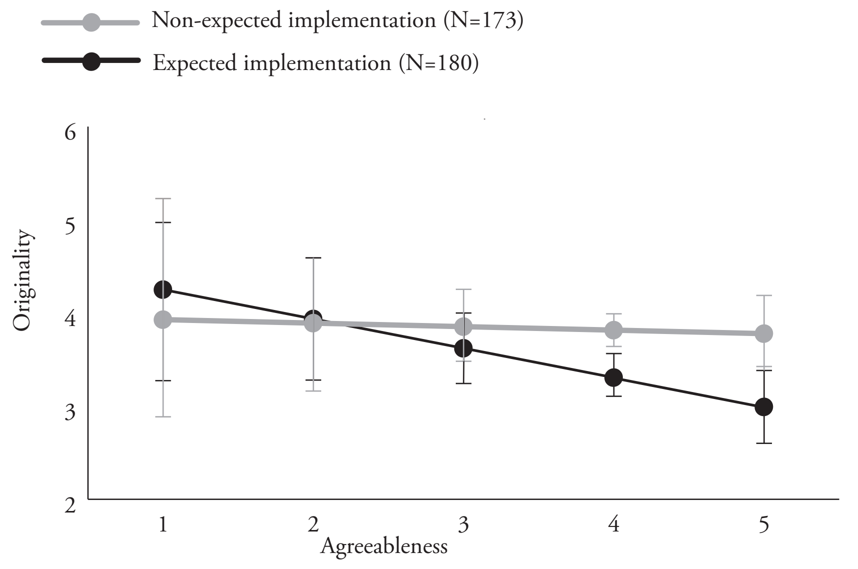


Hypotheses 3c predicted the effect of expected implementation on choosing feasible rather than original ideas to craft, would be stronger for children with high levels of agreeableness than children with low levels of agreeableness. Multivariate ordinary least squares regression suggests that the main effect of expected implementation on feasibility is significant (b = 2.33, $p < .05$, $\eta^2 = 0.01$), while the interaction effect with agreeableness is not (b = 0.04, $p > .10$; see Table 5.2). In addition, neither the main effect of expected implementation on originality is significant (b = -1.64, $p > .10$, $\eta^2 = 0.01$), nor the interaction effect with agreeableness (b = -0.28, $p > .10$; see Table 5.2). To illustrate the nature of this interaction, Figure 5.3c shows that the effect of expected implementation on idea selection in terms of feasibility is not different for children high or low on agreeableness. As such, hypothesis 3c is rejected.

Figure 5.3c

Interaction between expected implementation and agreeableness for (i) feasibility and (ii) originality





5.6 Discussion

The aim of this study was to examine whether children select less innovative ideas when they expect to implement these ideas on a later moment. More specifically, the present study aimed to investigate the effect of expected implementation of ideas on the selection of original and feasible ideas among primary school children. In addition, it was examined whether children's personality traits moderated this relationship.

As expected, the results demonstrated that children who expected implementation selected ideas that were less creative – i.e., more feasible and less original – than the ideas that were selected by children who did not had the expectation to later implement the selected ideas. This finding can be explained by the novelty and usefulness trade-off where original ideas are often perceived as useless or infeasible (Manske & Davis, 1968; Mueller et al., 2012; Rietzschel et al., 2006; Rietzschel et al., 2010; Ward, 2008). The results showed that children in grade 6 are aware that highly original ideas are often difficult to implement, as the feasibility and originality of their selected ideas where negatively correlated. This implies that the expectation of idea implementation, causes children to play it safe with regard to the practicality and to choose for the more feasible ideas rather than the more original ones.

In contrast, children who did not expect implementation were found to turn a blind eye to the feasibility aspects of ideas and focused on the originality aspect only. However, for an end product to be found creative, originality is not enough as it must also be feasible (Runco & Jaeger, 2012). Therefore, this study also investigated whether the level of feasibility of ideas would increase more than the level of originality would decrease for children who expected implementation - compared to children who did not expect implementation. This study indeed found that when children expected implementation, the level of feasibility of ideas decreased more than the level of originality did. This seems to imply that when children expect idea implementation, they become better able to juggle between the originality and feasibility aspects of ideas and select more creative ideas (that are both original and feasible) than children who did not expect such an implementation, and vice versa.

Furthermore, it was investigated whether children' personality traits play a moderating role in the relationship between expected implementation and idea selection. The results indicated that personality only played a minor role in creative idea selection. Only the personality trait conscientiousness was found to play a moderating role in idea selection. In the expected implementation condition, children with higher levels of conscientiousness were found to be stronger influenced to choose for feasibly rather than original ideas, than children demonstrating lower levels of conscientiousness. This finding can be explained by the fact that conscientious children tend to be attentive, persistent, orderly and neat, and think before acting (Shiner & Caspi, 2003). As such, it may be that even when they were instructed to select innovative ideas, they already took the possible implementation of these ideas in mind, and, consequently, put more focus on feasibility.

5.6.1 Implications for educational practice

The results of this study have important implications for educational practice. Our finding that children inhibited themselves in selecting original ideas once expecting idea implementation, suggests that this instruction approach may lead to the loss of potential novel ideas and out of fear that these may be impossible to implement. As such, primary schools teachers, who strive for the implementation of original ideas, are advised to not already mention the later implementation of the ideas when instructing children to select original ideas. Instead, teachers should support children when facing highly original, but seemingly unrealistic ideas, in thinking of ways how the idea can be made feasible. For instance, if a child struggles with an unfeasible idea, teachers could ask questions such as: 'what kind of other materials could you use to make it implementable?' or: 'just suppose that everything is possible, how would the idea become more feasible?' (see Seechaliao, 2017).

Another way to simulate children in choosing more creative ideas may be to remove the practical element of implementing ideas from the assignment. Focusing only on the end products runs the risk of overlooking the creative potential of children who may have had very original ideas but were too afraid to implement them. Consequently, a teacher may run the risk of applauding the creativity of a small group children who had the courage to try out their original ideas; while at the same time failing to recognize the creative potential of a group of equally creative children, who were too afraid to bring their ideas into practice. Focusing on the process and product may help educators draw out and support the development of children's creative potential (see Beghetto, 2010).

Furthermore, this study found that not all children were affected by expected implementation in their idea selection in the same way, as children demonstrating high levels of conscientiousness were less affected than children with low levels of conscientiousness. This difference between children is apparent in almost every classroom: some children are more conscientious than others. While less conscientious children tend to be easily distracted and more careless, conscientious children tend to be persistent, orderly and neat, and think before acting. This study found that conscientious children had a tendency to select more feasible ideas, even though they were instructed to select innovative ideas and did not expect idea implementation. Therefore, our recommendation would be to adapt the instruction of assignments to children's personality. For example, given conscientious children's preference for more feasible ideas, they would benefit from an instruction emphasizing the importance of original ideas as well.

5.6.2 Limitations and suggestions for future research

To the best of our knowledge, this is the first study demonstrating the effect of expected implementation on children's idea selection. Furthermore, it was investigated whether the relationship between expected implementation and idea selection differed for children with different personality traits. To test this, a relatively large sample of 403 children were asked in a randomized design to select two innovative ideas with or without the expectation to having to implement these ideas in the classroom.

While this study had many strengths, several limitations must also be acknowledged. First, this study investigated whether the expectation of idea implementation affects idea selection among primary school children. However, this study did not investigate the underlying mechanisms why this happens. Prior research has showed that people have a natural bias against creativity over feasibility because of uncertainty (e.g. Amabile, 1996; Mueller et al., 2012). As such, children may select different types of ideas to reduce uncertainty in the implementation phase. Future research should aim to investigate these underlying mechanisms.

Furthermore, this study has been conducted only among children in grade 6. Therefore this study's findings only apply to this age group. As such, the current study provides a starting point in the research on the effect of expected implementation on idea selection in children, but more research is needed to trace the development of the influence of expected implementation on idea selection in other age groups as well.

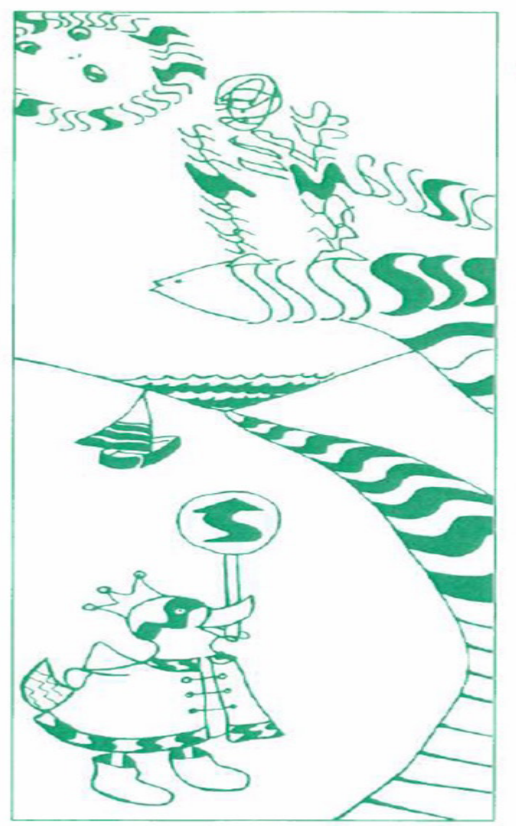
5.6.3 Conclusion

In sum, the present study investigated the effect of expected implementation of ideas on children's selection of original and feasible ideas. Furthermore, it was examined whether children's personality traits moderated this relationship. The results showed that expected implementation exerted different effects on the two dimension of final product innovation. Children who expected implementation selected less original ideas, but more feasible ideas than did children in the non-expected implementation condition. Furthermore, we found that the personality trait conscientiousness moderated this relationship. Children high on conscientiousness had a tendency to select more feasible even though they were instructed to select innovative ideas and did not expect idea implementation. The results highlight the importance for educators to carefully consider whether or not practical components should be part of assignments, and to tailor instruction in assignments to the individual needs of children. Together, this research provides a better understanding of the effects of expected implementation on idea selection

5.7 Appendix A

Online assignment for non-expected implementation condition (i.e., control group)

CREATIVE THINKING WITH WORDS



Name:

Number:

Continue to the next page when the teacher instructs you to continue.

Introduction assignment 1. Rating ideas:

In the middle of this page, you see a picture of a stuffed toy elephant. Ten children from primary school The Florest have generated ideas to improve a toy elephant so that children will have more fun playing with it.

For these generated ideas, how many children do you think came up with this idea? Give a rating **between** 1 and 10 for each idea. For example:

- If you think all children have generated the idea, you give the number 10 the idea (10 children came up with the idea).
- If you think 7 children have generated the idea, you give the number 7 to the idea (7 children came up with the idea).
- If you think 1 child has generated the idea, you give the number 1 to the idea (1 child came up with the idea).

Continue to the next page when the teacher instructs you to continue.

Practice assignment. Rating ideas:

Before we start, we will have a practice round. The goal of this practice round is to see whether everyone understands the assignment. You will receive three practice questions.

Practice question 1:

Out of 10 children, everyone has generated the following idea: *"Enlarge the elephant"*. What number do you fill in? ___

Practice question 2:

Out of 10 children, five children have generated the following idea: *"Make the elephant water resistant"*. What number do you fill in? ___

Practice question 3:

Out of 10 children, one child has generated the following idea: *"Make the elephant transparent"*. What number do you fill in? ___

Raise your hand when you are finished. The teacher will stop by.

Assignment 1. Read all ideas:

Now, you can read all ideas. We have combined ideas into one idea that were mentioned by several children.

Idea:	Implement magnets in the feet, so the toy elephant can be hung or can climb.
Idea:	Create a toy elephant that is able to move.
Idea:	Implement a zipper in the toy elephant so you can put stuff into it, like a puzzle, bag or surprise.
Idea:	Create a toy elephant that is able to poop.
Idea:	Create a robot elephant that you can control.
Idea:	Enlarge the toy elephant.
Idea:	Create a toy elephant that is able to drink water.
Idea:	Create a toy elephant that is water resistant.
Idea:	Create a toy elephant that is extra soft.
Idea:	Create toy elephants in various colors.
Idea:	Create a toy elephant that is able to talk or make sounds.
Idea:	Implement a mini-tablet on the elephant's belly with games.
Idea:	Attach unfolding tusks with sharp points to the toy elephant.
Idea:	Remove the tusks from the toy elephant.
Idea:	Implement balls in the feet, so that the toy elephant can roll.
Idea:	Attach bandage around the toy elephant, so that the elephant is sick or clumsy.
Idea:	Create a toy elephant that is able to fly.
Idea:	Create a toy elephant that is able to spit fire.
Idea:	Create a toy elephant from which you can brush the elephant's teeth.
Idea:	Create a toy elephant that is able to bounce.

Continue to the next page when you have read all ideas.

Assignment 2. Rating ideas:

Ten children from primary school The Florest have generated ideas to improve a toy elephant so that children will have more fun playing with it. We have combined ideas into one idea that were mentioned by several children.

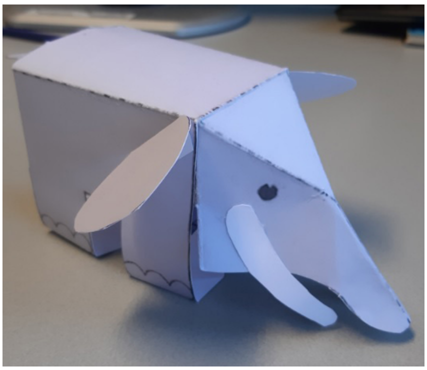
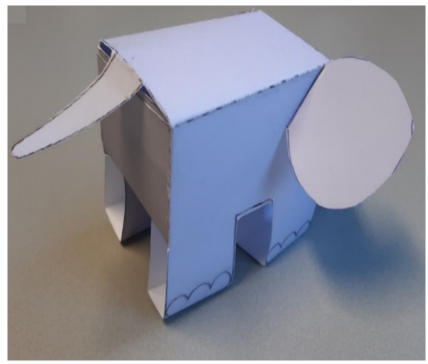
For the following ideas, how many children do you think came up with this idea?
Give a rating **between** 1 en 10 for each idea.
For example, if you think all children have generated the idea, you will give the number 10 the idea. If you think that 7 children have generated the idea, you will give the number 7 to the idea. If you think that 1 child has generated the idea, you will give the number 1 to the idea.

Idea:	Implement magnets in the feet, so the toy elephant can be hung or can climb.	________
Idea:	Create a toy elephant that is able to move.	________
Idea:	Implement a zipper in the toy elephant so you can put stuff into it, like a puzzle, bag or surprise.	________
Idea:	Create a toy elephant that is able to poop.	________
Idea:	Create a robot elephant that you can control.	________
Idea:	Enlarge the toy elephant.	________

Idea:	Create a toy elephant that is able to drink water.	________
Idea:	Create a toy elephant that is water resistant.	________
Idea:	Create a toy elephant that is extra soft.	________
Idea:	Create toy elephants in various colors.	________
Idea:	Create a toy elephant that is able to talk or make sounds.	________
Idea:	Implement a mini-tablet on the elephant's belly with games.	________
Idea:	Attach unfolding tusks with sharp points to the toy elephant.	________
Idea:	Remove the tusks from the toy elephant.	________
Idea:	Implement balls in the feet, so that the toy elephant can roll.	________
Idea:	Attach bandage around the toy elephant, so that the elephant is sick or clumsy.	________
Idea:	Create a toy elephant that is able to fly.	________
Idea:	Create a toy elephant that is able to spit fire.	________
Idea:	Create a toy elephant from which you can brush the elephant's teeth.	________
Idea:	Create a toy elephant that is able to bounce.	________

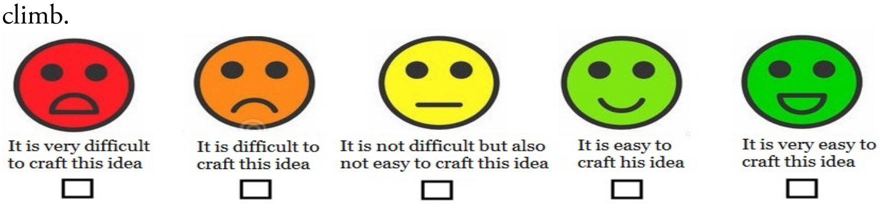
Continue to the next page.

Assignment 3. Rating ideas:
In the middle of this page, you see a picture of a paper toy elephant. Ten children from primary school The Florest have generated ideas to improve a toy elephant to that children will have more fun playing with it.

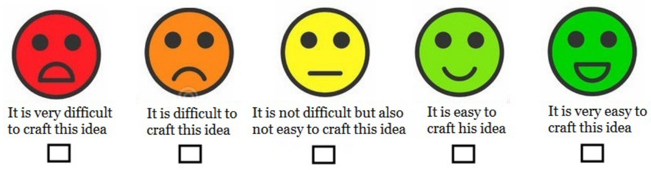
The children from primary school The Florest have been asked to craft an idea. They can use the following craft materials: pencil, pen, paper, scissors, glue, balls, magnets, iron wire/rope, paper clips, water, plastic tubes and bags, bouncy balls, bandages. The crafted toy elephant does not have to be really useful, but it should be clear what idea the child has crafted. The children will have approximately 50 minutes to craft one idea.
How difficult or easy do you think it would be to craft the following ideas? For each idea, tick one of the five faces that best shows how difficult or easy it seems to you to craft the idea.

Idea: Implement magnets in the feet, so the toy elephant can be hung or can climb.

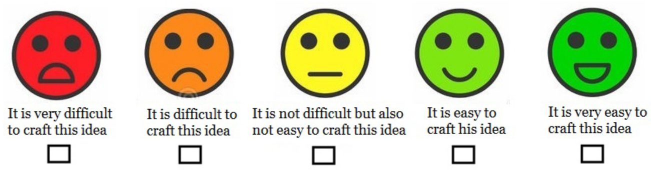


Continue to the next page.

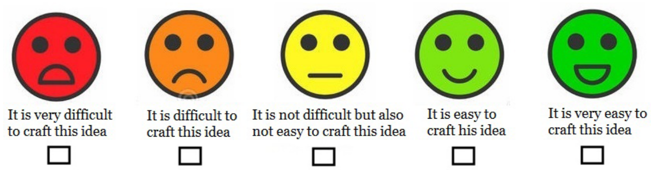
Idea: Create a toy elephant that is able to move.



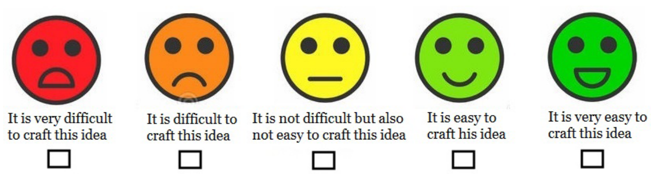
Idea: Implement a zipper in the toy elephant so you can put stuff into it, like a puzzle, bag or surprise.



Idea: Create a toy elephant that is able to poop.

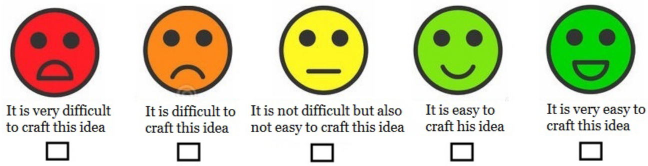


Idea: Create a toy elephant that you can control.

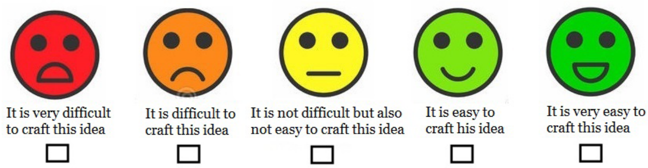


Continue to the next page.

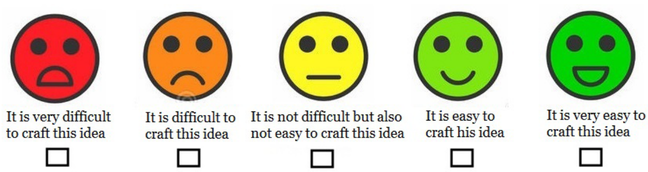
Idea: Enlarge the toy elephant.



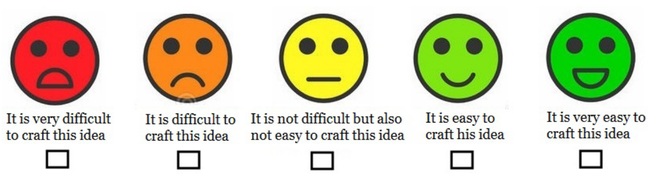
Idea: Create a toy elephant that is able to drink water.



Idea: Create a toy elephant that is water resistant.

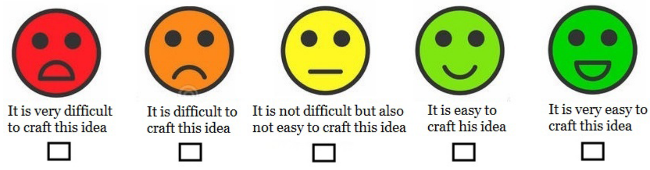


Idea: Create a toy elephant that is extra soft.

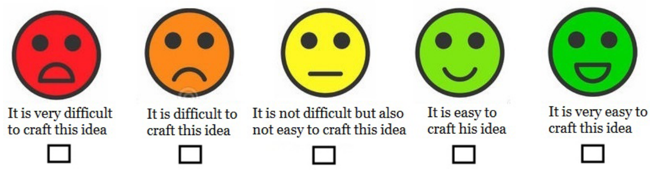


Continue to the next page.

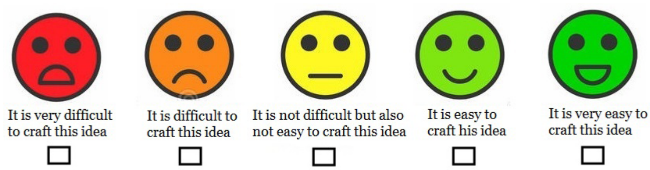
Idea: Create toy elephants in various colors.



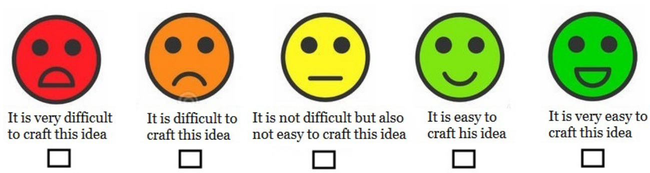
Idea: Create a toy elephant that is able to talk or make sounds.



Idea: Implement a mini-tablet on the elephant's belly with games.

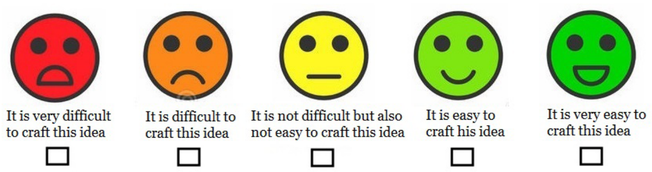


Idea: Attach unfolding tusks with sharp points to the toy elephant.

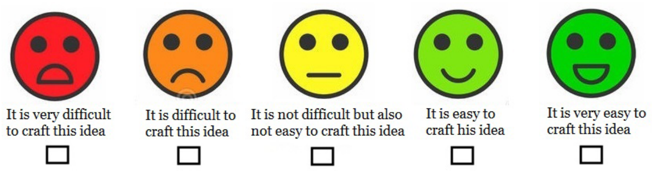


Continue to the next page.

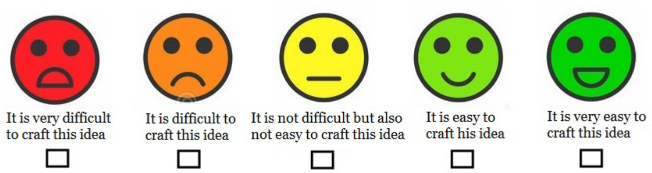
Idea: Remove the tusks from the toy elephant.



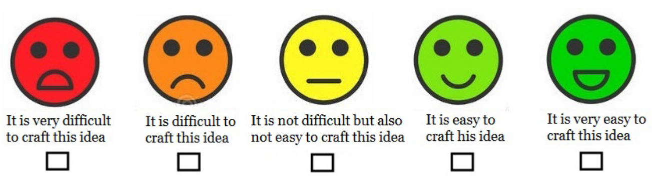
Idea: Implement balls in the feet, so that the toy elephant can roll.



Idea: Attach bandage around the toy elephant, so that the elephant is sick or clumsy.

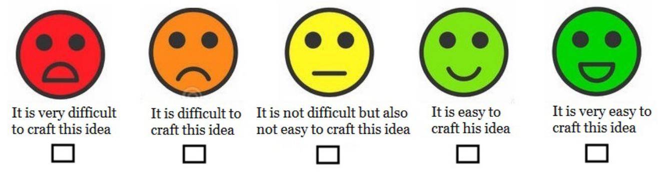


Idea: Create a toy elephant that is able to fly.



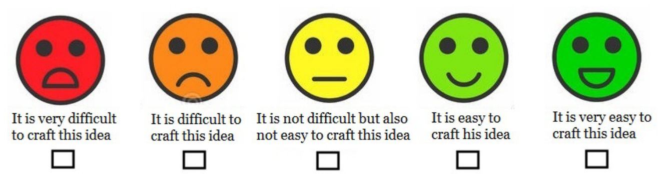
Continue to the next page.

Idea: Create a toy elephant that is able to spit fire.



Idea: Create a toy elephant from which you can brush the elephant’s teeth.

Idea: Create a toy elephant that is able to bounce.



Raise your hand and the teacher will stop by.

Assignment 4. Problem solving at the Zoo:

Like humans, monkeys like to play. When monkeys are bored in the Zoo, they exhibit negative behavior: they break things and bite each other. You will be asked to generate ideas for toys to prevent this problem. A condition for the toys is that the toy or part of the toy has to release food as a reward for the monkeys.

You will have 10 minutes to generate as many interesting and unusual ways as possible for toys for monkeys where food is relaesed as a reward. Write down one idea per line.

1.____

2.____

3.____

4.____

5.____

6.____

7.____

8.____

9.____

10._______________________________________

11._______________________________________

12._______________________________________

13._______________________________________

14.____

15.____

16.____

17.____

18.____

19.____

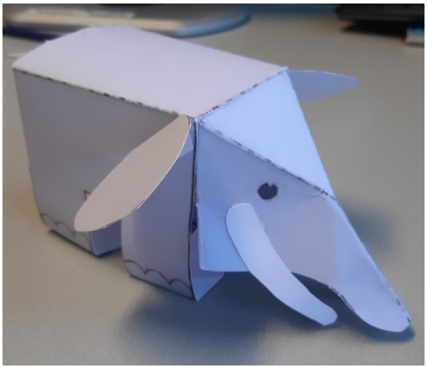
20.____

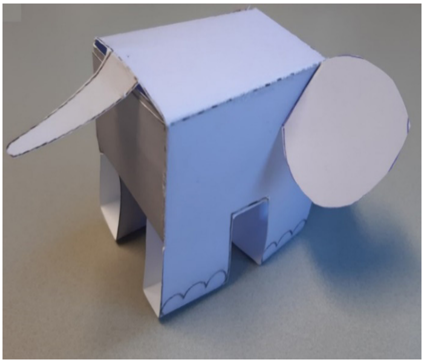
Continue to the next page.

Assignment 5. Choosing an innovative idea:

An **innovative idea** is a rare idea that you almost never see.

In the middle of this page, you see a picture of a paper toy elephant. Ten children from primary school The Florest have generated ideas to improve a toy elephant to that children will have more fun playing with it.





A toy factory needs your help! The toy factory makes toy animals, such as elephants, dogs, rabbits and so on. They would like to receive innovative ideas to change a toy elephant. They will first test these ideas on a toy elephant made of paper. **You will <u>NOT</u> craft these ideas, because you will be crafting ideas for monkey toys.**[34]

1. From the following ideas, mark the **<u>5 most innovative ideas</u>** (an innovative idea is a rare idea that you almost never see):

Idea:	Implement magnets in the feet, so the toy elephant can be hung or can climb.
Idea:	Create a toy elephant that is able to move.
Idea:	Implement a zipper in the toy elephant so you can put stuff into it, like a puzzle, bag or surprise.
Idea:	Create a toy elephant that is able to poop.
Idea:	Create a toy elephant that you can control.

34 This is the online assignment for the non-expected implementation condition (i.e., control group). The expected implementation condition received the same instructions, with exception of the manipulation: "A toy factory needs your help! The toy factory makes toy animals, such as elephants, dogs, rabbits and so on. They would like to receive innovative ideas to change a toy elephant. They will first test these ideas on a toy elephant made of paper. **You will craft these ideas.**"

Idea:	Enlarge the toy elephant.
Idea:	Create a toy elephant that is able to drink water.
Idea:	Create a toy elephant that is water resistant.
Idea:	Create a toy elephant that is extra soft.
Idea:	Create toy elephants in various colors.
Idea:	Create a toy elephant that is able to talk or make sounds.
Idea:	Implement a mini-tablet on the elephant's belly with games.
Idea:	Attach unfolding tusks with sharp points to the toy elephant.
Idea:	Remove the tusks from the toy elephant.
Idea:	Implement balls in the feet, so that the toy elephant can roll.
Idea:	Attach bandage around the toy elephant, so that the elephant is sick or clumsy.
Idea:	Create a toy elephant that is able to fly.
Idea:	Create a toy elephant that is able to spit fire.
Idea:	Create a toy elephant from which you can brush the elephant's teeth.
Idea:	Create a toy elephant that is able to bounce.

2. From your chosen 5 ideas, select the most **innovative idea** (an innovative idea is a rare idea that you almost never see):

The most innovative idea is:

____

Continue to the next page.

Assignment 6. Choosing an idea to craft:

In assignment 4, you have generated ideas for toys for monkeys.

The Zoo Artis needs your help! They would like to create a new toy for monkeys. You will be asked to select an idea for the Zoo Artis. After this assignment, you will craft this idea.

1. From all your ideas, select one idea to craft (1 answer):

I will craft the following idea: ______________________________________

Raise your hand and the teacher will stop by.

5.8 Appendix B

Table B.1

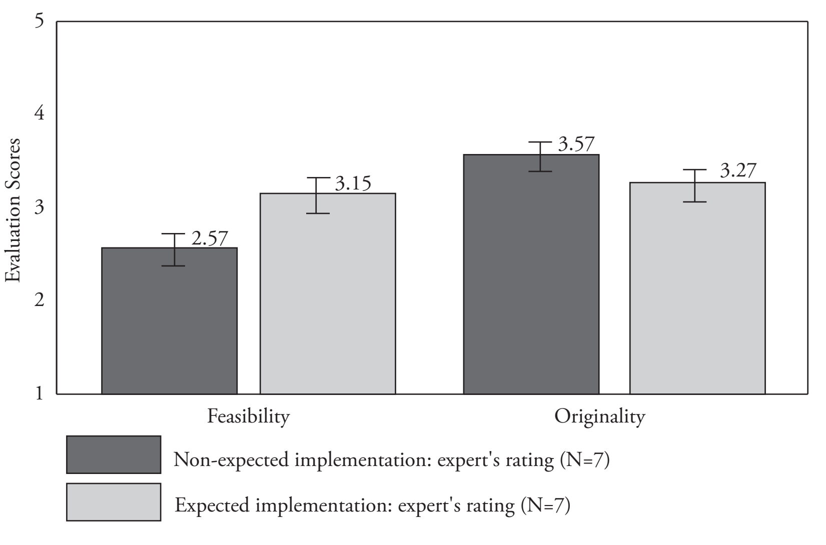
Ordinary Least Squares (OLS) regression with experts' average rating

	Feasibility Model 1a		Originality Model 1b	
	b-coefficient	η^2	*b*-coefficient	η^2
Constant	2.57 ***		4.14 ***	
	(0.05)		(0.05)	
Expected implementation	0.58 ***	0.13	-0.21 ***	0.02
	(0.08)		(0.07)	
R-squared	0.13		0.04	
F-statistics	59.43		8.41	
Number of observations	403		403	

Note: Standard errors in parentheses. Effect sizes were calculated as partial eta squares (η^2) where 0.01 is considered to be a small effect size, 0.09 medium effect size and 0.25 a large effect size. Significance levels indicated as follows: * $p < .05$, ** $p < .01$, *** $p < .001$.

Figure B.1.

Mean feasibility and originality of two selected ideas (experts' ratings)



5.9 Appendix C

Table C.1
Comparison Multivariate Ordinary Least Squares models and Conditional logit models (comparison Table 5.2 and Table 5.3)

	Feasibility		Originality	
	Multivariate OLS (*b*-coefficient)	Cconditional logit (odds ratio)	Multivariate OLS (*b*-coefficient)	Conditional logit (odds ratio)
Main effect				
Expected implementation	1.16 ***	2.03 ***	-0.43 ***	0.86 *
Interaction with personality				
Expected implementation	2.33 *	2.22 ***	-1.64	0.85 *
Personality traits				
Openness to experience	-0.08	0.93	0.09	1.07
Conscientiousness	0.37 **	1.26 *	-0.31 *	0.79 **
Extraversion	0.03	0.92	-0.03	0.95
Agreeableness	-0.04	0.98	-0.04	0.96
Emotional stability	0.10	1.02	-0.01	1.03
Interaction with personality				
*Implementation*openness*	-0.05	0.95	0.17	1.06
*Implementation*conscientiousness*	-0.54 **	0.74 **	0.38	1.28 *
*Implementation*extraversion*	0.17	1.15	0.08	1.02
*Implementation*agreeableness*	0.04	1.00	-0.28	0.97
*Implementation*emotional stability*	0.11	0.99	0.02	0.98

Note: In conditional logit, estimates greater than one are considered positive effects, while estimates smaller than one are considered negative effects.

6

Conclusion

6.1 Summary of main results and contributions

The aim of the thesis was to gain a deeper understanding of and improve the process of creative idea evaluation and idea selection among students. The 4 P's were used as a theoretical framework, because students' ability to *recognize* and *select* creative ideas (i.e., product) depends strongly on attitudes, dispositions, feelings, and beliefs (i.e., person), cognitive thinking processes (i.e., process), and is subject to situational constraints (i.e., press).

In Chapter 2, we contributed to the long-standing discussion whether creativity is domain-general or domain-specific by investigating creativity differences among university students. In case of domain-specificity, for example, a creative artist could differ from a creative engineer in terms of who they are (i.e., person), how they think (i.e., process), and even how they perceive creativity in artefacts (i.e., product). For this purpose, we examined these differences between (a) General Thematic Areas (Art and Science); (b) Specific *Science* domains (STEM), and; (c) Engineering micro-domains, for a total of 2277 German students. To this end, a series of one-way between groups multivariate analyses of covariance (MANCOVA) and a series of one-way analyses of covariance (ANCOVA) were conducted. The results showed many statistically significant, but uniformly small, differences at all levels, across a range of person, process and product variables. The pattern of results suggest that openness, creative self-efficacy and divergent thinking may be general pre-requisites for creativity. In contrast, the way that characteristics of creative products (i.e., originality, feasibility, effectiveness) are perceived seem to be more context dependent. These insights add to the hybrid approach literature that argues that creativity has both specific and general components (Baer & Kaufman, 2005; Kaufman & Baer, 2005; Plucker, Beghetto, & Dow, 2004).

In Chapter 3, we investigated whether students' creativity can be developed via a cognitive-based training programme. This 10-hour training incorporated two idea generation techniques (i.e., silent brainstorming, analogical thinking) and two idea evaluation techniques (i.e., idea evaluation metric, strengths and weakness analysis). A pre-post-test within-subject design was conducted among 51 Dutch undergraduate students. They participated in the training in the first or second educational semester. As such, students participated in both experimental conditions (control and intervention), albeit at different times (i.e., within subject design). General Linear Model (GLM) for repeated measures suggested that students generated more ideas (i.e., *fluency*) and different kind of ideas (i.e., *flexibility*) after training, but the results were non-significant. In line with prior research, the findings indicated that a cognitive-based training does not impact idea evaluation skills. This could stem from the fact that cognitive techniques - such as idea evaluation metric and strengths and weakness analysis – do not solely depend on domain knowledge, while idea evaluation is dependent on knowledge to assess novelty and feasibility of ideas and products (Cropley, 2006). As such, this chapter illustrates the importance of domain knowledge in the evaluation of ideas and adds to the emerging literature on the use and benefits of idea evaluation techniques (Vernon et al., 2016).

In Chapter 4, we investigated the effect of task exposure or familiarity on idea evaluation. Specifically, we investigated whether students become better able to recognize creative ideas from others, when they themselves earlier generated ideas for the same problem (as this provides them with more insight into the associative history of each ideas, and what ideas were rejected in favor of those that were kept). In many settings, people are responsible for evaluating ideas generated by others while they were not involved in the idea generation process, and, as such not exposed to the task. However, little is known on how this lack of task exposure affects idea evaluation. For this purpose, 1864 German students evaluated ideas on their creativity, originality and feasibility. Their ratings were compared to content experts' and creativity experts' ratings. The students were randomly assigned to one of the following conditions: task exposure (i.e., in this condition they had to generate and evaluate ideas for the same task) or no task exposure (i.e., in this condition they had to generate ideas for a different task than the idea evaluation task). The results showed that task exposure improves students' ability to accurately recognize creative and original ideas, and their ability to discriminate between highly feasible and unfeasible ideas. As such, these findings suggest that task exposure is beneficial for creative idea forecasting. Together, the results highlight the importance to carefully reconsider whether or not people should be exposed to the task before evaluating other's ideas.

In Chapter 5, we investigated the effect of expected implementation of ideas on children's selection of novel and feasible ideas. Worldwide, constructivist pedagogies have spread across primary schools that give children opportunity to develop 21st-century competences, such as creativity and innovation. An important characteristic of these pedagogies is that children are often asked to transform their ideas into tangible and physical products (Beghetto & Kaufman, 2010; Cardarello, 2014; Davies et al., 2013). However, it is unknown whether the expectation of having to implement an idea in practice affects the kind of ideas children select. For this purpose, 403 Dutch grade-6 children (age 10-13) selected two innovative ideas to improve the use of a stuffed toy elephant with or without the expectation to actually implement these ideas in the classroom. The results showed that children who expected implementation were less likely to select original ideas, but more likely to select feasible ideas than children who had no expectation to implement ideas. Moreover, implementation focused more on feasibility as compared to originality when selecting innovative ideas. The personality trait conscientiousness was found to moderate this relationship. Children with a high conscientiousness were found to select more feasible ideas even though they were instructed to select innovative ideas and did not expect idea implementation. These findings shed light on literature that implicitly assumed that idea implementation affects idea selection (Baer, 2012; Sharma, 1999). For instance, Sharma (1999) noted that many creative ideas are generated, but few reach the implementation phase.

6.2 Directions for future research

In sum, the chapters presented in this thesis show that the way how we think (i.e., *process*) and who we are (i.e., *person*) are general pre-requisites for creativity, while the way that creative *products* are recognized and selected is more context-specific. General cognitive techniques, such as strength and weaknesses analysis, seem not to benefit students' ability to recognize creative ideas, while task exposure does seem to improve students' idea evaluation. Even though students may be better at recognizing which ideas are creative, this thesis also shows that students inhibit themselves in selecting original ideas when they expect implementation of these ideas.

Notwithstanding the findings of this thesis, there are still many questions that would benefit from further academic research. Chapter 2 found that creativity has both general (i.e., process and person) and specific (i.e., product) components. Due to the cross-sectional nature of this study, we do not know whether this distinction emerged because of a selection from students in specific studies or developed by the study. Future research is needed to identify *when* these domain-general and domain-specific differences arise. For instance, future studies could employ a longitudinal design to address the question if a person is changed by what they study or whether different academic majors attract different kinds of people (i.e., selection). It should be noted as well that person, process and product factors of the 4P's framework were investigated, while we did not address *press* or the *environment* in which students operate. Future studies could further investigate differences in the environment or context, because we found that individual's assessment of product creativity are domain-specific.

In addition, the conclusion that general cognitive thinking does not develop idea evaluation skills has to be taken with caution. Only two techniques (i.e., idea evaluation metric and strength and weaknesses analysis) have been tested, while there are many more (e.g. repertory grids). As such, future research could investigate other types of idea evaluation techniques.

Moreover, there are several methodological approaches to assess idea evaluation accuracy. In this thesis, the widely accepted consensual assessment technique is employed where idea evaluation accuracy is assessed by the computing the average discrepancy between the participants' ratings and experts' ratings. Research has found that different types of experts have solid-to-strong inter-rater reliability when they rate the creativity of ideas or products (Amabile, 1996; Baer, Kaufman, & Gentile, 2004; Cheng, Wang, Liu, & Chen, 2010; Benedek et al., 2013). However, even though subjective ratings of creativity have been proven to be reliable, this method puts much weight on the experts' ratings as a golden standard criterion. As such, future research should employ various manners to assess accuracy (e.g. hit rates, discrepancy of evaluations with criterion scores, covariation of evaluation with criterion values). In addition, creativity consists of multiple criteria, such as originality and feasibility. As such, a natural question that remains unresolved is whether different types of experts agree on different criteria of creativity as well (e.g. originality and feasibility). Chapter 4 compared students' rating of ideas with that of *creativity* experts and *content* experts. Interestingly, while both types of experts agreed on the assessment of ideas in terms of their creativity and originality, they disagreed in their assessment of feasibility. As such, future studies may best consider experts with a certain degree of domain knowledge of the topic, when using expert ratings as method to evaluate participant ratings.

Furthermore, *good practices* such as power analyses, preregistration, making scientific papers and data publicly available (i.e., open access) has attracted growing interest in the field of economics and psychology.[35] Prior to the experiment in Chapter 5, a power analysis had been conducted to calculate the needed sample size given the expected effect size, alpha and power. However, no power-analysis had been conducted for Chapter 2 and 4. For Chapter 2, the data collection was part of a bi-annually survey-based research project "Fachkraft" that normally has sufficient power (N > 1500). For Chapter 4, the experiment was bound to the practicalities of

35 In preregistration, the research design, hypotheses, and analysis plan are specified before observing the outcomes of a study. Preregistration could improve research in two ways. First, preregistration provides a clear distinction between confirmatory research that uses data to test hypotheses and exploratory research that uses data to generate hypotheses. Second, journals favor submissions that report statistically significant effects, a bias that tends to inflate estimates of effect size in the published literature. Preregistering may reduce the influence of publication bias on effect-size estimation.

evaluating a training program in an existing bachelor program (N = 51). Moreover, while no one could be opposed to preregistration, the Chapters in this dissertation have not used preregistration, because this practice was relatively new and unknown to the researcher while collecting data (e.g. September 2017 and March 2018). Future studies would benefit from these good practices, because they greatly improve the quality of fundamental research.

Finally, this thesis found that students selected less original ideas, but more feasible ideas when they expected implementation of their ideas. In contrast, students without such an expectation selected more original ideas. Given that creative ideas are defined as a combination of original and feasible ideas, these results seems to suggest that students select more creative ideas when they expect implementation. However, students were not asked to select creative ideas, but original ideas. This was done, because the pilot study revealed that creativity as a concept is too difficult to grasp for children in grade 6. As such, it would be misleading to draw definitive conclusions about the selection of creative ideas. Future research could train children in understanding the concept of creativity to replicate this study.

7

Valorisation: implications for practice

7.1 Educational strategies

The results of Chapter 2 indicated that openness, creative self-efficacy and divergent thinking are general pre-requisites for creativity. This means that students in every subject matter benefit from a more open-minded attitude, more confidence in their own creative behavior and more experience in divergent thinking. As such, teachers do not have to formulate domain-specific strategies to nurture creativity (e.g. mathematical creativity). Instead, they can exploit any opportunity to foster creativity. For instance, a biology teacher might be responsible for courses in genetics, but also on ecology. While genetics is mostly about understanding DNA, ecology is much more about system thinking and less about specific components of the human body. Due to the general nature of creativity, the teacher can stimulate students' creative self-efficacy in both courses by providing them with positive feedback on their creative performance. In addition, it gives teachers the opportunity to decide which course assignments profit most from an open-ended structure to stimulate divergent thinking. For instance, there might be laboratory courses that do not lend themselves for open-ended assignments. In these courses, students simply need to learn the rules of the game, and, therefore, teachers could decide to formulate creativity-fostering strategies in other courses.

Related to course assignments, the findings of Chapter 5 suggest that the expectation of implementing ideas in practice inhibits students from selecting original ideas. If teachers find it important that students experiment with novel ideas and do not shy away from them, they may need to consider to leave the practical component out of the assignments. Another option would be to explicitly state that the level of novelty will be judged in the final product. The reward contingency literature has proven that individuals will act more creative when they are given contingent, and task-focused performance instructions (Byron & Khazanchi, 2012; Eisenberger & Armeli, 1997).

Next to assignments, a branch of research together with the findings of Chapter 3, suggest that cognitive techniques, such as silent brainstorming and analogical thinking, help students in generating more and different kind of ideas. As such, teachers can integrate idea generation techniques in their teaching strategies. For example, when students have to decide on a topic for their thesis, the teacher could ask students to individually generate ideas before having a traditional brainstorm with other classmates.

In addition, the results presented in Chapter 4 of this thesis suggest that simply asking teachers to evaluate creative ideas of students – without first generating their own creative ideas – may raise the probability of rejecting relative good ideas. The results indicated that teachers could become better in recognizing original and creative ideas from students if they are more involved in the generation process. As such, teachers could take a couple of minutes before rating assignments to generate their own ideas for the same problem. This small change in their teaching strategy gives them more task exposure, and this could provide them with more insight into the associative history of students' ideas, and what ideas were possible rejected in favor of those that were kept by students.

7.2 Recruitment of students

Next to teachers, the results of this thesis are highly relevant as well for those who are responsible for the recruitment of scarce talent for STEM (i.e., Science, Technology, Engineering and Mathematics). Chapter 2 shows that students in science, technology, engineering and mathematics possess similar personal traits and cognitive processes as art students (i.e., openness, creative self-efficacy and divergent thinking). Independent of academic subject matter, openness, creative self-efficacy and divergent thinking are general requisites of creativity. As such, practioners interested and responsible for attracting creative students can use these results in the recruitment process. For instance, they could highlight their support and mission for creativity and innovation in their information to potential future STEM-students. Another option may be to emphasize the openness as part of the study programme.

7.3 Design of curricula in education

As shown in Chapter 2, openness, creative self-efficacy and divergent thinking are important for creativity, independent of subject matter. As such, the fostering of openness, creative self-efficacy and divergent thinking should be fostered in every subject matter. For instance, not only assignments in STEM-based disciplines should offer open-ended assignments to develop divergent thinking, but also in business and economics, law, psychology and medicine.

Students in every subject matter benefit from a more open-minded attitude, more confidence in their own creative behavior and more experience in divergent thinking. In addition, chapter 3 suggest that cognitive techniques, such as brainstorming and analogical thinking, may help students to generate more and different kind of ideas (i.e., divergent thinking). As such, the development of creativity should be integrated in every part of the curricula and not only in specially designed courses or workshops about creativity. For example, an exam in economics tests students whether they understand the basic principles of economics. However, the learning approach to teach these principles could be done creatively (e.g. encourage students to identify and explore their knowledge, use cognitive techniques such as brainstorming and analogical thinking, and use imaginative approaches to make learning interesting and effective).

As shown in Chapter 4, teachers could become better able to recognize original and creative ideas of students when they first take the time to generate ideas for themselves. As such, we recommend to give teachers more time in rating assignments. Teachers can take this additional time to individually generate ideas before starting to rate assignments. The extra time can be a short duration of 4 minutes, for example, and could make all the difference in a creative idea from a student being rejected or recognized by teachers.

Curriculum Vitae

Kim obtained her bachelor degree in Sociology in 2014 at Radboud University. During her bachelors, she completed the honors programme. After graduating from her bachelor degree, she obtained her master degree in Sociology in 2015 at Radboud University. In September 2015, she joined the Research Centre for Education and Labour Market (ROA) at Maastricht University as a policy researcher. During her time as a policy researcher, Kim was involved in various educational- and labor market projects (e.g. labor market forecasts, national project to evaluate talent programs).

In September 2016, she started her PhD research on creativity and innovation in educational settings. The results of this PhD research are described in this thesis and were presented at several (inter)national conferences. In the final year of her PhD, Kim took the opportunity to work together with international expert David Cropley at the University of South Australia in Adelaide.

Next to her PhD, Kim obtained a university teacher qualification from the School of Business and Economics at Maastricht University in 2018. In this same year, she received an excellent grade from Maastricht's' Education Institute for her performance as a tutor. After finishing her PhD, Kim continued her line of research into the development of creativity and innovation in higher education and she is currently working as a postdoctoral researcher at Radboud University. Her research interests span from creativity and innovation in education to the measurement of creativity, and cover multiple creative stages such as idea generation, idea evaluation and idea selection.

Bibliography

Abraham, A., Asquith, S., Ahmed, H., & Bourisly, A. K. (2019). Comparing the efficacy of four brief inductions in boosting short-term creativity. *Journal of Cognitive Enhancement, 3*(1), 85-93.

Acar, S., & Runco, M. A. (2012). Creative abilities: Divergent thinking. In *Handbook of organizational creativity* (pp. 115-139): Elsevier.

Agogué, M., Le Masson, P., Dalmasso, C., Houdé, O., & Cassotti, M. (2015). Resisting classical solutions: The creative mind of industrial designers and engineers. *Psychology of Aesthetics, Creativity, the Arts, 9*(3), 313.

Allen, J., Belfi, B., & Borghans, L. (2020). Is there a rise in the importance of socioemotional skills in the labour market? Evidence from a trend study among college graduates. *Frontiers in Psychology, 11*, 1710.

Amabile, T.M. (1982). Children's artistic creativity: Detrimental effects of competition in a field setting. *Personality and Social Psychology Bulletin, 8*(3), 573-578.

Amabile, T.M. (1984). Social inuences on creativity: Evaluation, coaction, and surveillance. *Annual Convention of the American Psychological Association*, 92, 24-28.

Amabile, T.M., Hennessey, B. A., & Grossman, B. S. (1986). Social influences on creativity: The effects of contracted-for reward. *Journal of Personality and Social Psychology, 50*(1), 14-23.

Amabile, T.M., Hennessey, B.A., & Mueller, J.S. (1999). Consensual assessment. *Encyclopedia of creativity*, 1, 347-359.

Amabile, T.M. & Pratt, M. G. (2016). The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in Organizational Behavior, 36*, 157-183.

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology, 43*(5), 997.

Amabile, T. M. (1996). Creativity in Context. Boulder, Colorado. In: Westview Press Inc.

Anders Ericsson, K., Roring, R. W., & Nandagopal, K. (2007). Giftedness and evidence for reproducibly superior performance: An account based on the expert performance framework. *High ability studies*, 18(1), 3-56.

Anwar, M. N., & Aness, M. (2012). An examination of the relationship between creative thinking and academic achievements of secondary school students. *International Interdisciplinary Journal of Education, 1*(219), 1-4.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics, 118*(4), 1279-1333.

Baas, M., De Dreu, C. K., & Nijstad, B. A. (2011). Creative production by angry people peaks early on, decreases over time, and is relatively unstructured. *Journal of Experimental Social Psychology, 47*(6), 1107-1115.

Baepler, P., Walker, J., & Driessen, M. (2014). It's not about seat time: Blending, flipping, and efficiency in active learning classrooms. *Computers & Education, 78*, 227-236.

Baer, J., & Kaufman, J. C. (2005). Bridging generality and specificity: The amusement park theoretical (apt) model of creativity. *Roeper Review, 27*(3), 158-163.

Baer, J., Kaufman, J. C., & Gentile, C. A. (2004). Extension of the consensual assessment technique to nonparallel creative products. *Creativity research journal*, 16(1), 113-117.

Baer, M. (2012). Putting creativity to work: The implementation of creative ideas in organizations. *Academy of Management Journal, 55*(5), 1102-1119.

Bakhshi, H., Downing, J. M., Osborne, M. A., & Schneider, P. (2017). *The future of skills: Employment in 2030*: Pearson.

Basadur, M., Runco, M. A., & Vega, L. A. (2000). Understanding how creative thinking skills, attitudes and behaviors work together: A causal process model. *The Journal of Creative Behavior, 34*(2), 77-100.

Batey, M., Chamorro-Premuzic, T., & Furnham, A. (2009). Intelligence and personality as predictors of divergent thinking: The role of general, fluid and crystallised intelligence. *Thinking Skills and Creativity, 4*(1), 60-69.

Batey, M., & Furnham, A. (2006). Creativity, intelligence, and personality: A critical review of the scattered literature. *Genetic, Social, and General Psychology Monographs, 132*(4), 355-429.

Batey, M., & Furnham, A. (2008). The relationship between measures of creativity and schizotypy. *Personality and Individual Differences, 45*(8), 816-821.

Beggan, J. K. (1992). On the social nature of nonsocial perception: The mere ownership effect. *Journal of Personality and Social Psychology, 62*(2), 229-237.

Beggan, J. K., & Brown, E. M. (1994). Association as a psychological justification for ownership. *The journal of psychology, 128*(4), 365-380.

Beghetto, R. A. (2006). Creative self-efficacy: Correlates in middle and secondary students. *Creativity Research Journal, 18*(4), 447-457.

Beghetto, R. A., & Kaufman, J. C. (2010). *Nurturing creativity in the classroom*: Cambridge University Press.

Benedek, M., Mühlmann, C., Jauk, E., & Neubauer, A. C. (2013). Assessment of divergent thinking by means of the subjective top-scoring method: Effects of the number of top-ideas and time-on-task on reliability and validity. *Psychology of Aesthetics, Creativity, and the Arts, 7*(4), 341.

Benedek, M., Nordtvedt, N., Jauk, E., Koschmieder, C., Pretsch, J., Krammer, G., & Neubauer, A. (2016). Assessment of creativity evaluation skills: A psychometric investigation in prospective teachers. *Thinking Skills and Creativity, 21*, 75-84.

Berg, J. M. (2016). Balancing on the creative highwire: Forecasting the success of novel ideas in organizations. *Administrative Science Quarterly, 61*(3), 433-468.

Birdi, K. (2007). A lighthouse in the desert? Evaluating the effects of creativity training on employee innovation. *Journal of Creative Behavior, 41*(4), 249-270.

Birdi, K. (2016). Creativity training. In *Human resource management, innovation and performance* (pp. 298-312): Springer.

Blair, C. S., & Mumford, M. D. (2007). Errors in idea evaluation: Preference for the unoriginal? *The Journal of Creative Behavior, 41*(3), 197-222.

Bødker, S., & Iversen, O. S. (2002). *Staging a professional participatory design practice: moving PD beyond the initial fascination of user involvement.* Paper presented at the Proceedings of the second Nordic conference on Human-computer interaction.

Brawner, C. E., Camacho, M. M., Lord, S. M., Long, R. A., & Ohland, M. W. (2012). Women in industrial engineering: Stereotypes, persistence, and perspectives. *Journal of Engineering Education, 101*(2), 288-318.

Bruch, M. A., & Krieshok, T. S. (1981). Investigative versus realistic Holland types and adjustment in theoretical engineering majors. *Journal of Vocational Behavior, 18*(2), 162-173.

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of personality and social psychology*, 67(3), 366.

Byrge, C., & Hansen, S. (2013). Course in new thinking in higher education: Enhancing creativity through the means of training, theory and workshop. *Problems of Education in the 21st Century, 51*(51), 18-32.

Byron, K., & Khazanchi, S. (2012). Rewards and creative performance: a meta-analytic test of theoretically derived hypotheses. *Psychological Bulletin, 138*(4), 809.

Cachia, R., Ferrari, A., Ala-Mutka, K., & Punie, Y. (2010). Creative learning and innovative teaching: Final report on the study on creativity and innovation in education in EU member states. *JRC Working Papers*, (JRC62370).

Cardarello, R. (2014). *Enhancing scientific thinking in children: Suggestions based on studies about creativity.* Paper presented at the Conference proceedings. New perspectives in science education.

Castillo, L. C. (1998). The effect of analogy instruction on young children's metaphor comprehension. *Roeper Review, 21*(1), 27-31.

Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). *Handbook for the 16PF test*: IPAT: Champaign.

Charles, R. E., & Runco, M. A. (2001). Developmental trends in the evaluative and divergent thinking of children. *Creativity Research Journal, 13*(3-4), 417-437.

Cheng, Y.-Y., Wang, W.-C., Liu, K.-S., & Chen, Y.-L. (2010). Effects of association instruction on fourth graders' poetic creativity in Taiwan. *Creativity Research Journal, 22*(2), 228-235.

Choi, J. N. (2004). Individual and contextual predictors of creative performance: The mediating role of psychological processes. *Creativity Research Journal, 16*(2-3), 187-199.

Choi, J. N. (2007). Change-oriented organizational citizenship behavior: effects of work environment characteristics and intervening psychological processes. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior, 28*(4), 467-484.

Cohen, J. (1988). Statistical power analysis for the behavioural sciences, 2nd edn.(Hillsdale, NJ: L. Erlbaum Associates).

Corazza, G. E. (2016). Potential originality and effectiveness: The dynamic definition of creativity. *Creativity Research Journal, 28*(3), 258-267.

Costa, P. T., & McCrea, R. R. (1992). *Revised neo personality inventory (neo pi-r) and neo five-factor inventory (neo-ffi)*: Psychological Assessment Resources.

Cropley, A. (2006). In praise of convergent thinking. *Creativity Research Journal, 18*(3), 391-404.

Cropley, D. H., & Cropley, A. (2005). Engineering creativity: A systems concept of functional creativity. *Creativity across domains: Faces of the muse, 1*(15), 169-185.

Cropley, D. H. (2015). Promoting creativity and innovation in engineering education. *Psychology of Aesthetics, Creativity, and the Arts, 9*(2), 161.

Cropley, D. H. (2019). *Creativity, innovation and the future of engineering.* Paper presented at the World Engineers Convention 2019, Melbourne.

Cropley, D. H., Cropley, A., & Sandwith, B. L. (2017). *Creativity in the Engineering Domain*. New York, NY: Cambridge University Press.

Cropley, D. H., & Kaufman, J. C. (2012). Measuring functional creativity: Non-expert raters and the Creative Solution Diagnosis Scale. *The Journal of Creative Behavior, 46*(2), 119-137.

Cropley, D. H., Kaufman, J. C., & Cropley, A. (2011). Measuring creativity for innovation management. *Journal of technology management & innovation, 6*(3), 13-30.

Cseh, G. M., & Jeffries, K. K. (2019). A scattered CAT: A critical evaluation of the consensual assessment technique for creativity research. *Psychology of Aesthetics, Creativity, and the Arts, 13*(2), 159.

Curtin, D. W. (1982). The aesthetic dimension of science: 1980 Nobel conference. *New York: Philosophical Library*.

Dahl, D. W., & Moreau, P. (2002). The influence and value of analogical thinking during new product ideation. *Journal of Marketing Research, 39*(1), 47-60.

Dailey, L., & Mumford, M. D. (2006). Evaluative aspects of creative thought: Errors in appraising the implications of new ideas. *Creativity Research Journal, 18*(3), 385-390.

Daly, S. R., Christian, J. L., Yilmaz, S., Seifert, C. M., & Gonzalez, R. (2012). Assessing design heuristics for idea generation in an introductory engineering course. *International Journal of Engineering Education, 28*(2), 463.

Dane, E. (2010). Reconsidering the trade-off between expertise and flexibility: A cognitive entrenchment perspective. *Academy of Management Review, 35*(4), 579-603.

Davies, D., Jindal-Snape, D., Collier, C., Digby, R., Hay, P., & Howe, A. (2013). Creative learning environments in education—A systematic literature review. *Thinking Skills and Creativity, 8*, 80-91.

de Buisonjé, D. R., Ritter, S. M., de Bruin, S., ter Horst, J. M., & Meeldijk, A. (2017). Facilitating creative idea selection: The combined effects of self-affirmation, promotion focus and positive affect. *Creativity Research Journal, 29*(2), 174-181.

Dean, D. L., Hender, J. M., Rodgers, T. L., & Santanen, E. (2006). Identifying good ideas: constructs and scales for idea evaluation. *Journal of the Association for Information Systems, 7*(10), 646-699.

Denker, F., Eling, K., & Herstatt, C. (2016, July 3-6). *The Role Of Individuals' Domain Knowledge In Evaluating Radical New Product Ideas*. [Conference presentation]. R&D Management Conference 2016 "From Science to Society: Innovation and Value Creation", Cambridge, United Kingdom.

Dexter, S., & Kozbelt, A. (2013). Free and open source software (FOSS) as a model domain for answering big questions about creativity. *Mind & Society, 12*(1), 113-123.

Diedrich, J., Benedek, M., Jauk, E., & Neubauer, A. C. (2015). Are creative ideas novel and useful? *Psychology of Aesthetics, Creativity, and the Arts, 9*(1), 35.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*(1), 417-440.

Dyson, S. B., Chang, Y.-L., Chen, H.-C., Hsiung, H.-Y., Tseng, C.-C., & Chang, J.-H. (2016). The effect of tabletop role-playing games on the creative potential and emotional creativity of Taiwanese college students. *Thinking Skills and Creativity, 19*, 88-96.

Edwards, M., & McGoldrick, C. (2006). Creativity and curricula in higher education: academics' perspectives. In *Developing creativity in higher education* (pp. 79-93): Routledge.

Eisenberger, R., & Armeli, S. (1997). Can salient reward increase creative performance without reducing intrinsic creative interest? *Journal of Personality and Social Psychology, 72*(3), 652.

Eisenman, R. (1969). Creativity and academic major: Business versus English majors. *Journal of Applied Psychology, 53*(5), 392.

Eling, K., Griffin, A., & Langerak, F. (2016). Consistency Matters in Formally Selecting Incremental and Radical New Product Ideas for Advancement. *Journal of product innovation management, 33*(S1), 20-33.

Elsbach, K. D., & Kramer, R. M. (2003). Assessing creativity in Hollywood pitch meetings: Evidence for a dual-process model of creativity judgments. *Academy of Management Journal, 46*(3), 283-301.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, 41(4), 1149-1160.

Faure, C. (2004). Beyond brainstorming: Effects of different group procedures on selection of ideas and satisfaction with the process. *The Journal of Creative Behavior, 38*(1), 13-34.

Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review, 2*(4), 290-309.

Fogler, H. S., LeBlanc, S. E., & Rizzo, B. R. (1995). *Strategies for creative problem solving*: PTR Prentice Hall Englewood Cliffs, NJ.

Förster, J. (2009). Cognitive consequences of novelty and familiarity: How mere exposure influences level of construal. *Journal of Experimental Social Psychology, 45*(2), 444-447.

Frederiksen, M. H., & Knudsen, M. P. (2017). From creative ideas to innovation performance: The role of assessment criteria. *Creativity and Innovation Management, 26*(1), 60-74.

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change, 114*, 254-280.

Furnham, A., Batey, M., Booth, T. W., Patel, V., & Lozinskaya, D. (2011). Individual difference predictors of creativity in Art and Science students. *Thinking Skills and Creativity, 6*(2), 114-121.

Furnham, A., & Crump, J. (2013). The sensitive, imaginative, articulate art student and conservative, cool, numerate science student: Individual differences in art and science students. *Learning and Individual Differences, 25*, 150-155.

Fürst, G., Ghisletta, P., & Lubart, T. (2016). Toward an Integrative Model of Creativity and Personality: Theoretical Suggestions and Preliminary Empirical Testing. *Journal of Creative Behavior, 50*(2), 87-108.

Gassmann, O., & Zeschky, M. (2008). Opening up the solution space: the role of analogical thinking for breakthrough product innovation. *Creativity and Innovation Management, 17*(2), 97-106.

George, J. M., & Zhou, J. (2001). When openness to experience and conscientiousness are related to creative behavior: an interactional approach. *Journal of Applied Psychology, 86*(3), 513.

Goldberg, L. R. (1990). An alternative" description of personality": the big-five factor structure. *Journal of Personality and Social Psychology, 59*(6), 1216.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*(1), 84-96.

Goldstone, R. L., Medin, D. L., & Halberstadt, J. (1997). Similarity in context. *Memory & Cognition*, 25(2), 237-255.

Greene, D., & Lepper, M. R. (1974). Effects of extrinsic rewards on children's subsequent intrinsic interest. *Child Development*, 1141-1145.

Grohman, M., Wodniecka, Z., & Kłusak, M. (2006). Divergent thinking and evaluation skills: Do they always go together? *The Journal of Creative Behavior, 40*(2), 125-145.

Grosul, M., & Feist, G. J. (2014). The creative person in science. *Psychology of Aesthetics, Creativity, and the Arts, 8*(1), 30.

Guilford, J. P. (1959). Three faces of intellect. *American Psychologist, 14*(8), 469-479.

Guilford, J. P. (1967). *The nature of human intelligence*. New York, NY: McGraw-Hill

Guo, J., Ge, Y., & Pang, W. (2019). The underlying cognitive mechanisms of the rater effect in creativity assessment: The mediating role of perceived semantic distance. *Thinking Skills and Creativity, 33*, 100572.

Haddad, C., & Roman, M. (2004). FRED SMITH: OVERNIGHT SENSATION. *Business Week*(3865), 42-42.

Hartlet, J., & Greggs, M. A. (1997). Divergent Thinking in Arts and Science Students. *Studies in Higher Education, 22*(1), 93-97.

Hawksworth, J., Berriman, R., & Goel, S. (2018). Will robots really steal our jobs? An international analysis of the potential long term impact of automation. *PricewaterhouseCoopers, http://pwc.co.uk/economics, access, 13*.

Hennessey, B. A. (1994). The consensual assessment technique: An examination of the relationship between ratings of product and process creativity. *Creativity Research Journal, 7*(2), 193-208.

Herman, A., & Reiter-Palmon, R. (2011). The effect of regulatory focus on idea generation and idea evaluation. *Psychology of Aesthetics, Creativity, and the Arts, 5*(1), 13-20.

Hommel, B., Colzato, L. S., Fischer, R., & Christoffels, I. (2011). Bilingualism and creativity: Benefits in convergent thinking come with losses in divergent thinking. *Frontiers in Psychology*, 2(273), 1-5.

Huber, J. C. (1998). Invention and inventivity as a special kind of creativity, with implications for general creativity. *The Journal of Creative Behavior*, 32(1), 58-72.

Hyypiä, M., & Parjanen, S. (2013). Boosting creativity with transformational leadership in fuzzy front end innovation processes. *Interdisciplinary Journal of Information, Knowledge and Management*, 8, 21-41.

IBM. (2010). *Capitalizing on complexity: Insights from the global chief executive officer study*: IBM.

Illies, J. J., & Reiter-Palmon, R. (2004). The Effects of Type and Level of Personal Involvement on Information Search and Problem Solving. *Journal of Applied Social Psychology*, 34, 1-22.

Im, H., Hokanson, B., & Johnson, K. K. P. (2015). Teaching creative thinking skills: A longitudinal study. *Clothing and Textiles Research Journal, 33*(2), 129-142.

International Labor Organization. (2017). *World employment and social outlook: Trends 2017.* Geneva: ILO.

Isaksen, S. G., Dorval, K. B., & Treffinger, D. J. (2010). *Creative approaches to problem solving: A framework for innovation and change*. California, CA: Sage Publications.

Jackson, R. L., & Hogg, M. A. (2010). Self-Enhancement Theory. In *Encyclopedia of Identity* (pp. 691-694). Thousand Oaks SAGE Publications, Inc. .

Jastrzębska, D., & Limont, W. (2017). Not only jumps, slumps, but also mini plateau. Creative potential assessed by the test for creative thinking-drawing production. A cross-sectional study of polish students aged from 7 to 18. *Creativity Research Journal, 29*(3), 337-342.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research, 2*(1999), 102-138.

Johnson, H. M. (1911). Clever hans (the horse of mr. von osten): A contribution to experimental, animal, and human psychology. *The Journal of Philosophy, Psychology and Scientific Methods, 8*(24), 663-666.

Jordan, T. (2013, November 1). No winter sun? Norwergian mountain town installs giant mirrors. *CNN*. Retrieved from Retrieved from https://edition.cnn.com/travel/article/rjukan-sun-mirror

Kapoor, H., & Khan, A. (2019). Creativity in context: Presses and task effects in negative creativity. *Psychology of Aesthetics, Creativity, and the Arts, 13*(3), 314-321.

Kaufman, J. C., & Baer, J. (2005). *Creativity across domains: Faces of the muse*: Psychology Press

Kaufman, J. C., Baer, J., Cole, J. C., & Sexton, J. D. (2008). A comparison of expert and nonexpert raters using the consensual assessment technique. *Creativity Research Journal*, 20(2), 171-178.

Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four C model of creativity. *Review of General Psychology, 13*(1), 1.

Kaufman, J. C., Niu, W., Sexton, J. D., & Cole, J. C. (2010). In the eye of the beholder: Differences across ethnicity and gender in evaluating creative work. *Journal of Applied Social Psychology, 40*(2), 496-511.

Kaufman, J. C., Baer, J., Cropley, D. H., Reiter-Palmon, R., & Sinnett, S. (2013). Furious activity vs. understanding: How much expertise is needed to evaluate creative work?. *Psychology of Aesthetics, Creativity, and the Arts*, 7(4), 332.

Kaufman, J. C., Pumaccahua, T. T., & Holt, R. E. (2013a). Personality and creativity in realistic, investigative, artistic, social, and enterprising college majors. *Personality Individual Differences, 54*(8), 913-917.

Kaufman, J. C., Pumaccahua, T. T., & Holt, R. E. (2013b). Personality and creativity in realistic, investigative, artistic, social, and enterprising college majors. *Personality and Individual Differences, 54*(8), 913-917.

Kaufman, J. C., Quilty, L. C., Grazioplene, R. G., Hirsh, J. B., Gray, J. R., Peterson, J. B., & DeYoung, C. G. (2016). Openness to experience and intellect differentially predict creative achievement in the arts and sciences. *Journal of Personality, 84*(2), 248-258.

Kaufman, J. C., & Sternberg, R. J. (2019). Cambridge handbook of creativity. In. New York, NY: Cambridge University Press.

Khatena, J. (1971). Teaching disadvantaged preschool children to think creatively with pictures. *Journal of Educational Psychology, 62*(5), 384.

Kim, K. H. (2011). The creativity crisis: The decrease in creative thinking scores on the Torrance Tests of Creative Thinking. *Creativity Research Journal, 23*(4), 285-295.

Kim, S., Chung, K., & Yu, H. (2013). Enhancing digital fluency through a training program for creative problem solving using computer programming. *The Journal of Creative Behavior, 47*(3), 171-199.

Kimbler, D. (1995). The development of modern industrial engineering. *Engineering horizons*.

Kline, P., & Lapham, S. L. (1992). Personality and faculty in British universities. *Personality and Individual Differences, 13*(7), 855-857.

Kozbelt, A., Dexter, S., Dolese, M., & Seidel, A. (2012). The aesthetics of software code: A quantitative exploration. *Psychology of Aesthetics, Creativity, and the Arts, 6*(1), 57.

Kunreuther, H., Novemsky, N., & Kahneman, D. (2001). Making low probabilities useful. *Journal of Risk and Uncertainty, 23*(2), 103-120.

Langholtz, H., Gettys, C., & Foote, B. (1995). Are resource fluctuations anticipated in resource allocation tasks? *Organizational Behavior and Human Decision Processes, 64*(3), 274-282.

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. J*ournal of Personality and social Psychology*, 28(1), 129-137.

Lewandowsky, S., Little, D., & Kalish, M. L. (2007). Knowledge and expertise. *Handbook of applied cognition, 830109*, 125-155.

Licuanan, B. F., Dailey, L. R., & Mumford, M. D. (2007). Idea evaluation: Error in evaluating highly original ideas. *The Journal of Creative Behavior, 41*(1), 1-27.

Lievens, F., Coetsier, P., De Fruyt, F., & De Maeseneer, J. (2002). Medical students' personality characteristics and academic performance: A five-factor model perspective. *Medical Education, 36*(11), 1050-1056.

Lim, H. S., & Choi, J. N. (2009). Testing an alternative relationship between individual and contextual predictors of creative performance. *Social Behavior and Personality: an international journal, 37*(1), 117-135.

Litchfield, R. C., Gilson, L. L., & Gilson, P. W. (2015). Defining creative ideas: Toward a more nuanced approach. *Group & Organization Management, 40*(2), 238-265.

Liu, H. Y., Wang, I. T., Chen, N. H., & Chao, C. Y. (2020). Effect of creativity training on teaching for creativity for nursing faculty in Taiwan: A quasi-experimental study. *Nurse education today*, 85, 1-7.

Lonergan, D. C., Scott, G. M., & Mumford, M. D. (2004). Evaluative aspects of creative thought: Effects of appraisal and revision standards. *Creativity Research Journal, 16*(2-3), 231-246.

Lounsbury, J. W., Smith, R. M., Levy, J. J., Leong, F. T., & Gibson, L. W. (2009). Personality characteristics of business majors as defined by the big five and narrow personality traits. *Journal of Education for Business, 84*(4), 200-205.

Loveland, K. K., & Olley, J. G. (1979). The effect of external reward on interest and quality of task performance in children of high and low intrinsic motivation. *Child Development*, 1207-1210.

Lubart, T. I., & Guignard, J. (2004). The generality-specificity of creativity: A multivariate approach. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity from potential to realization* (pp. 43–56). Washington, DC: American Psychological Association.

Lucas, B., Claxton, G., & Spencer, E. (2013). *Progression in student creativity in school: First steps towards new forms of formative assessment* (OECD Education Working Papers, No. 86). Paris, France: OECD Publishing.

Magnusson, P. R., Netz, J., & Wästlund, E. (2014). Exploring holistic intuitive idea screening in the light of formal criteria. *Technovation, 34*(5-6), 315-326.

Mainemelis, C. (2010). Stealing fire: Creative deviance in the evolution of new ideas. *Academy of Management Review, 35*(4), 558-578.

Manske, M. E., & Davis, G. A. (1968). Effects of simple instructional biases upon performance in the unusual uses test. *The Journal of general psychology, 79*(1), 25-33.

Marrs, H., Barb, M.R., Ruggiero, J.C. (2007). Self-Reported Influences on psychology major choices and personality. *Individual Differences Research*, 5(4), 289-299.

Mathisen, G. E., & Bronnick, K. S. (2009). Creative self-efficacy: An intervention study. *International Journal of Educational Research, 48*(1), 21-29.

Matraeva, A. D., Rybakova, M. V., Vinichenko, M. V., Oseev, A. A., & Ljapunova, N. V. (2020). Development of Creativity of Students in Higher Educational Institutions: Assessment of Students and Experts. *Universal Journal of Educational Research, 8*(1), 8-16.

McCarthy, M., Chen, C. C., & McNamee, R. C. (2018). Novelty and Usefulness Trade-Off: Cultural Cognitive Differences and Creative Idea Evaluation. *Journal of Cross-Cultural Psychology, 49*(2), 171-198.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics, 3*(4), 303-328.

McGraw, K. O. (1978). The detrimental effects of reward on performance: A literature review and a prediction model. *The hidden costs of reward: New perspectives on the psychology of human motivation*, 33-60.

McLoughlin, C., & Lee, M. J. (2008). The three p's of pedagogy for the networked society: Personalization, participation, and productivity. *International Journal of Teaching and Learning in Higher Education, 20*(1), 10-27.

Mednick, S. (1962). The associative basis of the creative process. *Psychological Review, 69*(3), 220.

Meinel, M., Wagner, T. F., Baccarella, C. V., & Voigt, K. I. (2019). Exploring the effects of creativity training on creative performance and creative self-efficacy: Evidence from a longitudinal study. *The Journal of Creative Behavior*, 53(4), 546-558.

Miron-Spektor, E., & Beenen, G. (2015). Motivating creativity: The effects of sequential and simultaneous learning and performance achievement goals on product novelty and usefulness. *Organizational Behavior and Human Decision Processes, 127*, 53-65.

Miron-Spektor, E., Gino, F., & Argote, L. (2011). Paradoxical frames and creative sparks: Enhancing individual creativity through conflict and integration. *Organizational Behavior and Human Decision Processes, 116*(2), 229-240.

Mount, M. K., & Barrick, M. R. (1995). The Big Five personality dimensions: Implications for research and practice in human resources management. *Research in personnel and human resources management, 13*(3), 153-200.

Mudge, J. F., Baker, L. F., Edge, C. B., & Houlahan, J. E. (2012). Setting an optimal α that minimizes errors in null hypothesis significance tests. *PloS one*, 7(2), e32734.

Mueller, J. S., Melwani, S., & Goncalo, J. A. (2012). The bias against creativity: Why people desire but reject creative ideas. *Psychological Science, 23*(1), 13-17.

Mumford, M. D., Baughman, W. A., & Sager, C. E. (2003). Picking the right material: Cognitive processing skills and their role in creative thought. In M. A. Runco (Ed.), *Critical and creative thinking* (pp. 19-68). Cresskill, NJ: Hampton.

Mumford, M. D., Lonergan, D. C., & Scott, G. (2002). Evaluating creative ideas: Processes, standards, and context. *Inquiry: Critical thinking across the disciplines, 22*(1), 21-30.

Mumford, M. D., Mobley, M. I., Reiter-Palmon, R., Uhlman, C. E., & Doares, L. M. (1991). Process analytic models of creative capacities. *Creativity Research Journal, 4*(2), 91-122.

Myszkowski, N., & Storme, M. (2019). Judge response theory? A call to upgrade our psychometrical account of creativity judgments. *Psychology of Aesthetics, Creativity, and the Arts*, 13(2), 167–175.

Nazzal, L. J. (2015). *Engineering creativity: Differences in creative problem solving stages across domains* (Doctoral dissertation). Retrieved from https://opencommons.uconn.edu/dissertations/753

Nemeth, C. J. (1997). Managing innovation: When less is more. *California Management Review, 40*(1), 59-74.

Nijstad, B. A., De Dreu, C. K., Rietzschel, E. F., & Baas, M. (2010). The dual pathway to creativity model: Creative ideation as a function of flexibility and persistence. *European Review of Social Psychology, 21*(1), 34-77.

Nijstad, B. A., & Stroebe, W. (2006). How the group affects the mind: A cognitive model of idea generation in groups. *Personality and Social Psychology Review, 10*(3), 186-213.

Nikander, J. B., Liikkanen, L. A., & Laakso, M. (2014). The preference effect in design concept evaluation. *Design Studies, 35*(5), 473-499.

Oldham, G. R., & Cummings, A. (1996). Employee creativity: Personal and contextual factors at work. *Academy of Management Journal, 39*(3), 607-634.

Petervari, J., Osman, M., & Bhattacharya, J. (2016). The Role of Intuition in the Generation and Evaluation Stages of Creativity. *Frontiers in Psychology*, 7, 1-12.

Piagèt, J. (1964). Development and learning. *International Journal of Teaching and Learning in Higher Education, 40*(2), 176-186.

Plucker, J. A., & Beghetto, R. A. (2004). *Why creativity is domain general, why it looks domain specific, and why the distinction does not matter*. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 153-167).

Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational Psychologist, 39*(2), 83-96.

Porter, A. C., & Raudenbush, S. W. (1987). Analysis of covariance: Its model and use in psychological research. *Journal of Counseling Psychology, 34*(4), 383.

Potočnik, K., & Anderson, N. (2016). A constructively critical review of change and innovation-related concepts: towards conceptual and operational clarity. *European Journal of Work and Organizational Psychology, 25*(4), 481-494.

Pringle, C. D., DuBose, P. B., & Yankey, M. D. (2010). Personality characteristics and choice of academic major: Are traditional stereotypes obsolete? *College Student Journal, 44*(1), 131-143.

Puccio, G. J. (2017). From the dawn of humanity to the 21st century: Creativity as an enduring survival skill. *The Journal of Creative Behavior, 51*(4), 330-334.

Putman, V. L., & Paulus, P. B. (2009). Brainstorming, brainstorming rules and decision making. *The Journal of Creative Behavior, 43*(1), 29-40.

Ratcheva, V., Leopold, T. A., & Zahidi, S. (2020). J*obs of tomorrow: mapping opportunity in the new economy*. In World Economic Forum, Geneva, Switzerland.

Reichardt, C. S. (1979). The statistical analysis of data from nonequivalent group designs. *Quasi-experimentation: Design and analysis issues for field settings*, 147-205.

Reilly, R. R., Lynn, G. S., & Aronson, Z. H. (2002). The role of personality in new product development team performance. *Journal of Engineering and Technology Management, 19*(1), 39-58.

Reiter-Palmon, R., Mumford, M. D., O'Connor Boes, J., & Runco, M. A. (1997). Problem construction and creativity: The role of ability, cue consistency, and active processing. *Creativity Research Journal*, 10(1), 9-23.

Rhodes, M. (1961). An analysis of creativity. *The Phi Delta Kappan, 42*(7), 305-310.

Rietzschel, E. F., Nijstad, B. A., & Stroebe, W. (2006). Productivity is not enough: A comparison of interactive and nominal brainstorming groups on idea generation and selection. *Journal of Experimental Social Psychology*, 42(2), 244-251.

Rietzschel, E. F., Nijstad, B. A., & Stroebe, W. (2010). The selection of creative ideas after individual idea generation: Choosing between creativity and impact. *British journal of psychology*, 101(1), 47-68.

Rietzschel, E. F., Nijstad, B. A., & Stroebe, W. (2007). Relative accessibility of domain knowledge and creativity: The effects of knowledge activation on the quantity and originality of generated ideas. *Journal of Experimental Social Psychology, 43*(6), 933-946.

Ritter, S. M., & Mostert, N. (2017). Enhancement of creative thinking skills using a cognitive-based creativity training. *Journal of Cognitive Enhancement*, 1(3), 243-253.

Ritter, S. M., Gu, X., Crijns, M., & Biekens, P. (2020). Fostering students' creative thinking skills by means of a one-year creativity training program. *PloS one*, 15(3), e0229773.

Ritter, S. M., van Baaren, R. B., & Dijksterhuis, A. (2012). Creativity: The role of unconscious processes in idea generation and idea selection. *Thinking Skills and Creativity, 7*(1), 21-27.

Rodriguez, W. A., Cheban, Y., Shah, S., & Watts, L. L. (2020). The general factor of personality and creativity: Diverging effects on intrapersonal and interpersonal idea evaluation. *Personality and Individual Differences, 167*, 110229.

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity research journal*, 24(1), 92-96.

Runco, M. A., & Smith, W. R. (1992). Interpersonal and intrapersonal evaluations of creative ideas. *Personality and Individual Differences*, 13(3), 295-302.

Runco, M. A., & Basadur, M. (1993). Assessing ideational and evaluative skills and creative styles and attitudes. *Creativity and Innovation Management, 2*(3), 166-173.

Runco, M. A., & Dow, G. T. (2004). Assessing the accuracy of judgments of originality on three divergent thinking tests. *The International Journal of Creativity & Problem Solving, 14*(2), 5-14.

Runco, M. A., Illies, J. J., & Eisenman, R. (2005). Creativity, originality, and appropriateness: What do explicit instructions tell us about their relationships? *The Journal of Creative Behavior, 39*(2), 137-148.

Russ, S. W., & Fiorelli, J. A. (2010). Developmental approaches to creativity. *The Cambridge handbook of creativity*, 233-249.

Sagone, E., & De Caroli, M. E. (2012). Creativity and thinking styles in arts, sciences, and humanities high school students. *International Journal of Developmental and Educational Psychology*, 1(1), 441-449.

Said-Metwaly, S., Fernández-Castilla, B., Kyndt, E., Van den Noortgate, W., & Barbot, B. (2020). Does the Fourth-Grade Slump in Creativity Actually Exist? A Meta-analysis of the Development of Divergent Thinking in School-Age Children and Adolescents. *Educational Psychology Review*, 33, 275-298.

Sarkar, P., & Chakrabarti, A. (2008). The effect of representation of triggers on design outcomes. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 22*(2), 101-116.

Scott, G., Leritz, L. E., & Mumford, M. D. (2004). The effectiveness of creativity training: A quantitative review. *Creativity Research Journal, 16*(4), 361-388.

Scott, G., Leritz, L. E., & Mumford, M. D. (2004b). Types of creativity training: Approaches and their effectiveness. *Journal of Creative Behavior, 38*(3), 149-179.

Seegers, P. K., Bergerhoff, J., Hartmann, S., & Knappe, A. (2019). *Fachkraft 2030: Erhebung zur wirtschaftlichen und allgemeinen Lebenssituation der Studierenden in Deutschland*. Köln, Germany.

Shalley, C. E., & Perry-Smith, J. E. (2001). Effects of social-psychological factors on creative performance: The role of informational and controlling expected evaluation and modeling experience. *Organizational Behavior and Human Decision Processes, 84*(1), 1-22.

Shalley, C. E., & Zhou, J. (2008). Organizational creativity research: A historical overview. *Handbook of organizational creativity*, 331, 3-31.

Sharma, A. (1999). Central dilemmas of managing innovation in large firms. *California Management Review, 41*(3), 146-164.

Shiner, R., & Caspi, A. (2003). Personality differences in childhood and adolescence: Measurement, development, and consequences. *Journal of child psychology and psychiatry*, 44(1), 2-32.

Silvia, P. J. (2008). Discernment and creativity: How well can people identify their most creative ideas? *Psychology of Aesthetics, Creativity, and the Arts, 2*(3), 139-146.

Simonton, D. K. (2003). Scientific creativity as constrained stochastic behavior: the integration of product, person, and process perspectives. *Psychological Bulletin, 129*(4), 475.

Simonton, D. K. (2004). *Creativity in science: Chance, logic, genius, and zeitgeist*: Cambridge University Press.

Simonton, D. K. (2009). Varieties of (scientific) creativity: A hierarchical model of domain-specific disposition, development, and achievement. *Perspectives on Psychological Science, 4*(5), 441-452.

Snow, C. P. (1959). Two cultures. *Science, 130*(3373), 419-419.

Stam, D., de Vet, A., Barkema, H. G., & De Dreu, C. K. (2013). Suspending group debate and developing concepts. *Journal of product innovation management, 30*, 48-61.

Sternberg, R. J., & Lubart, T. I. (1999). *The concept of creativity: Prospects and paradigms* (Vol. 1). Cambridge, England: Cambridge University press.

Stierand, M., & Dörfler, V. (2016). The role of intuition in the creative process of expert chefs. *The Journal of Creative Behavior, 50*(3), 178-185.

Stroebe, W., Nijstad, B. A., & Rietzschel, E. F. (2010). Beyond productivity loss in brainstorming groups: The evolution of a question. In *Advances in Experimental Social Psychology* (Vol. 43, pp. 157-203): Elsevier.

Sun, M., Wang, M., & Wegerif, R. (2019). Using computer-based cognitive mapping to improve students' divergent thinking for creativity development. *British Journal of Educational Technology, 50*(5), 2217-2233.

Tauber, A. I. (1996). *The Elusive Synthesis: Aesthetics and Science: Aesthetics and Science* (Vol. 182): Springer Science & Business Media.

Tesauro, G. (1989). *Connectionist learning of expert preferences by comparison training.* Paper presented at the Advances in Neural Information Processing Systems.

Tierney, P., & Farmer, S. M. (2002). Creative self-efficacy: Its potential antecedents and relationship to creative performance. *Academy of Management Journal, 45*(6), 1137-1148.

Toh, C. A., & Miller, S. R. (2016a). Choosing creativity: the role of individual risk and ambiguity aversion on creative concept selection in engineering design. *Research in Engineering Design, 27*(3), 195-219.

Toh, C. A., & Miller, S. R. (2016b). Creativity in design teams: the influence of personality traits and risk attitudes on creative concept selection. *Research in Engineering Design, 27*(1), 73-89.

Torrance, E. (1974). *The Torrance tests of creative thinking-TTCT Manual and Scoring Guide: Verbal test A, figural test*. Lexington, KY: Ginn.

Treffinger, D. J. (2007). Creative Problem Solving (CPS): Powerful tools for managing change and developing talent. *Gifted and Talented International, 22*(2), 8-18.

Tsakanikos, E., & Claridge, G. (2005). More words, less words: verbal fluency as a function of 'positive'and 'negative'schizotypy. *Personality and Individual Differences, 39*(4), 705-713.

Tubb, A. L., Cropley, D. H., Marrone, R. L., Patston, T., & Kaufman, J. C. (2020). The development of mathematical creativity across high school: Increasing, decreasing, or both? *Thinking Skills and Creativity, 35*, 100634.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*(4), 327.

Valgeirsdottir, D., & Onarheim, B. (2017). Studying creativity training programs: A methodological analysis. *Creativity and Innovation Management, 26*(4), 430-439.

Vally, Z., Salloum, L., AlQedra, D., El Shazly, S., Albloshi, M., Alsheraifi, S., & Alkaabi, A. (2019). Examining the effects of creativity training on creative production, creative self-efficacy, and neuro-executive functioning. *Thinking Skills and Creativity*, 31, 70-78.

van Breukelen, G. J. (2013). ANCOVA versus CHANGE from baseline in nonrandomized studies: The difference. *Multivariate Behavioral Research, 48*(6), 895-922.

van Broekhoven, K., Cropley, D.H., & Seegers, P. (2020). Differences in creativity across Art and STEM students: We are more alike than unalike. *Thinking Skills and Creativity*, 38, 1-13.

Van Damme, M., Anseel, F., Duyck, W., & Rietzschel, E. F. (2019). Strategies to improve selection of creative ideas: An experimental test of epistemic and social motivation in groups. *Creativity and Innovation Management, 28*(1), 61-71.

Vernon, D., Hocking, I., & Tyler, T. C. (2016). An evidence-based review of creative problem solving tools. *Human Resource Development Review, 15*(2), 230-259.

Veurink, N., & Sorby, S. (2013). Comparison of spatial skills of students entering different engineering majors. *The Engineering Design Graphics Journal, 76*(3).

Veurink, N. L., & Hamlin, A. (2011). *Spatial Visualization Skills: Impact on Confidence in an Engineering Curriculum.* Paper presented at the American Society for Engineering Education.

Vincent-Lancrin, S., González-Sancho, C., Bouckaert, M., Fernández-Barrerra, M., Jacotin, G., Urgel, J., & Vidal, Q. (Eds.). (2019). *Fostering Students' Creativity and Critical Thinking: What it Means in School.* Educational Research and Innovation. OECD Publishing: Paris.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*: Harvard university press.

Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior, 17*(1), 3-28.

Wallach, M. A., & Kogan, N. (1965), *Modes of Thinking in Young Children*, New York: Holt, Rinehart & Winston.

Wang, J., Wang, L., Liu, R.-D., & Dong, H.-Z. (2017). How expected evaluation influences creativity: Regulatory focus as moderator. *Motivation and Emotion, 41*(2), 147-157.

Ward, T. B. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology, 27*(1), 1-40.

Ward, T. B. (2008). The role of domain knowledge in creative generation. *Learning and Individual Differences, 18*(4), 363-366.

Wechsler, J. (1979). *Aesthetics in Science*. Cambridge, MA: MIT Press.

West, M., Farr, J., West, M., & Farr, J. (1990). Innovation at work, Innovation creativity at work. *Psychological and organizational strategies*, 3-13.

Wiley, J. (1998). Expertise as mental set: The effects of domain knowledge in creative problem solving. *Memory & Cognition, 26*(4), 716-730.

World Economic Forum (2016). *The future of jobs: Employment, skills and workforce strategy for the fourth industrial revolution.* Paper presented at the Global Challenge Insight Report, World Economic Forum, Geneva.

World Economic Forum (2018). *The future of jobs report 2018*.

Yuan, F., & Zhou, J. (2008). Differential effects of expected external evaluation on different parts of the creative idea production process and on final product creativity. *Creativity Research Journal, 20*(4), 391-403.

Yuan, R., & Zhou, J. (2002). *Expected evaluations and creativity: New answers to an old question.* Paper presented at the Annual Conference of the Academy of Management, Denver, CO.

Zacher, H., Robinson, A. J., & Rosing, K. (2016). Ambidextrous Leadership and Employees' Self-reported Innovative Performance: The Role of Exploration and Exploitation Behaviors. *Journal of Creative Behavior, 50*(1), 24-46.

Zare, B. (2011). *Creativity differences between art and Engineering students.* Paper presented at the Proceedings of the International Conference on E-business, Management and Economics.

Zeng, L., Salvendy, G., & Zhang, M. (2009). Factor structure of web site creativity. *Computers in Human Behavior, 25*(2), 568-577.

Zhou, C. (2012). Fostering creative engineers: A key to face the complexity of engineering practice. *European Journal of Engineering Education, 37*(4), 343-353.

Zhou, J., & Oldham, G. R. (2001). Enhancing creative performance: Effects of expected developmental assessment strategies and creative personality. *The Journal of Creative Behavior, 35*(3), 151-167.

Zhu, Y., Ritter, S. M., & Dijksterhuis, A. (2020). Creativity: intrapersonal and interpersonal selection of creative ideas. *The Journal of Creative Behavior*, 54(3), 626-635.

ROA Dissertation Series

1. Lex Borghans (1993), *Educational Choice and Labour Market Information*, Maastricht, Research Centre for Education and the Labour Market.
2. Frank Cörvers (1999), *The Impact of Human Capital on International Competitiveness and Trade Performance of Manufacturing Sectors*, Maastricht, Research Centre for Education and the Labour Market.
3. Ben Kriechel (2003), *Heterogeneity Among Displaced Workers*, Maastricht, Research Centre for Education and the Labour Market.
4. Arnaud Dupuy (2004), *Assignment and Substitution in the Labour Market*, Maastricht, Research Centre for Education and the Labour Market.
5. Wendy Smits (2005), *The Quality of Apprenticeship Training, Conflicting Interests of Firms and Apprentices*, Maastricht, Research Centre for Education and the Labour Market.
6. Judith Semeijn (2005), *Academic Competences and Labour Market Entry: Studies Among Dutch Graduates*, Maastricht, Research Centre for Education and the Labour Market.
7. Jasper van Loo (2005), *Training, Labour Market Outcomes and Self-Management*, Maastricht, Research Centre for Education and the Labour Market.
8. Christoph Meng (2005), *Discipline-Specific or Academic? Acquisition, Role and Value of Higher Education Competencies*, Maastricht, Research Centre for Education and the Labour Market.
9. Andreas Ammermüller (2007), *Institutional Effects in the Production of Education: Evidence from European Schooling Systems*, Maastricht, Research Centre for Education and the Labour Market.
10. Bart Golsteyn (2007), *The Ability to Invest in Human Capital*, Maastricht, Research Centre for Education and the Labour Market.
11. Raymond Montizaan (2010), *Pension Rights, human capital development and well-being*, Maastricht, Research Centre for Education and the Labour Market.
12. Annemarie Nelen (2012), *Part-Time Employment and Human Capital Development*, Maastricht, Research Centre for Education and the Labour Market.
13. Jan Sauermann (2013), *Human Capital, Incentives, and Performance Out- comes*, Maastricht, Research Centre for Education and the Labour Market

14. Harald Ulrich Pfeifer (2013), *Empirical Investigations of Costs and Benefits of Vocational Education and Training*, Maastricht, Research Centre for Education and the Labour Market.
15. Charlotte Büchner (2013), *Social Background, Educational Attainment and Labor Market Integration: An Exploration of Underlying Processes and Dynamics*, Maastricht, Research Centre for Education and the Labour Market.
16. Martin Humburg (2014), *Skills and the Employability of University Graduates*, Maastricht, Research Centre for Education and the Labour Market.
17. Jan Feld (2014), *Making the Invisible Visible, Essays on Overconfidence, Discrimination and Peer Effects*, Maastricht, Research Centre for Education and the Labour Market.
18. Olga Skriabikova (2014), *Preferences, Institutions, and Economic Outcomes: an Empirical Investigation*, Maastricht, Research Centre for Education and the Labour Market.
19. Gabriele Marconi (2015), *Higher Education in the National and Global Economy*, Maastricht, Research Centre for Education and the Labour Market.
20. Nicolas Salamanca Acosta (2015), *Economic Preferences and Financial Risk Taking*, Maastricht, Research Centre for Education and the Labour Market.
21. Ahmed Elsayed Mohamed (2015), *Essays on Working Hours*, Maastricht, Research Centre for Education and the Labour Market.
22. Roxanne Amanda Korthals (2015), *Tracking Students in Secondary Education, Consequences for Student Performance and Inequality*, Maastricht, Research Centre for Education and the Labour Market.
23. Maria Zumbuehl (2015), *Economic Preferences and Attitudes: Origins, Behavioral Impact, Stability and Measurement*, Maastricht, Research Centre for Education and the Labour Market.
24. Anika Jansen (2016), *Firms' incentives to provide apprenticeships – Studies on expected short- and long-term benefits*, Maastricht, Research Centre for Education and the Labour Market.
25. Jos Maarten Arnold Frank Sanders (2016), *Sustaining the employability of the low skilled worker: Development, mobility and work redesign*, Maastricht, Research Centre for Education and the Labour Market.
26. Marion Collewet (2017), *Working hours: preferences, well-being and productivity*, Maastricht, Research Centre for Education and the Labour Market.
27. Tom Stolp (2018), *Sorting in the Labor Market: The Role of Risk Preference and Stress*, Maastricht, Research Centre for Education and the Labour Market.

28. Frauke Meyer (2019), *Individual motives for (re-) distribution*, Maasricht, Research Centre for Education and the Labour Market.
29. Maria Ferreira Sequeda (2019), *Human Capital Development at School and Work*, Maasricht, Research Centre for Education and the Labour Market.
30. Marie-Christine Martha Fregin (2019), *Skill Matching and Outcomes: New Cross-Country Evidence*, Maasricht, Research Centre for Education and the Labour Market.
31. Sanne Johanna Leontien van Wetten (2020), *Human capital and employee entrepreneurship: The role of skills, personality characteristics and the work context*, Maasricht, Research Centre for Education and the Labour Market.
32. Cécile Alice Jeanne Magnée (2020), *Playing the hand you're dealt, The effects of family structure on children's personality and the effects of educational policy on educational outcomes of migrant children*, Maastricht, Research Centre for Education and the Labour Market.
33. Merve Nezihe Özer (2020), *Essays on Drivers and Long-Term Impact of Migration*, Maastricht, Research Centre for Education and the Labour Market.
34. Inge Ingeborg Henrica Maria Hooijen (2021), *Place attractiveness, A study of the determinants playing a role in residential settlement behaviour*, Maastricht, Research Centre for Education and the Labour Market.
35. Alexandra Marie Catherine de Gendre (2021), *Behavioral Barriers to Success in Education*, Maastricht, Research Centre for Education and the Labour Market.
36. Kim Christine Maria van Broekhoven (2021), *From creativity to innovation: Understanding and improving the evaluation and selection of ideas in educational settings*, Maastricht, Research Centre for Education and the Labour Market.