# FROM QUANTITY TO QUALITY

## Cognitive, motivational and social aspects of creative idea generation and selection

# FROM QUANTITY TO QUALITY

## Cognitive, motivational and social aspects of creative idea generation and selection

Van kwantiteit naar kwaliteit:
Cognitieve, motivationele en sociale aspecten
van het genereren en selecteren van creatieve ideeën
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de Rector Magnificus,
Prof. Dr. W. H. Gispen
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen op
dinsdag 6 december 2005, des middags te 12.45 uur

door

Eric Fulco Rietzschel
geboren op 23 september 1973, te Delft

Promotoren: Prof. Dr. W. Stroebe (Universiteit Utrecht)
Prof. Dr. C. K. W. de Dreu (Universiteit van Amsterdam)

Co-promotor: Dr. B. A. Nijstad (Universiteit van Amsterdam)

This project was subsidized by the Netherlands Organization for Scientific Research (NWO), NWO grant 410-21-0.

"Who is this guy? So he makes a study. I couldn't care less."

Red Auerbach

"In fact," he went on, holding his head down, and his voice getting lower and lower, "I don't believe that pudding ever *was* cooked! In fact, I don't believe that pudding ever *will* be cooked! And yet it was a very clever pudding to invent."

Lewis Carroll

# Preface

In 2001, I applied for a position as a PhD student at the University of Amsterdam, on a project concerning group creativity. I had no particular views on that topic, but thought it would be an interesting and stimulating field to do research in. One thing that particularly attracted me was the prospect of studying actual *behavior*. In the previous year, I had received my Master's degree in Psychonomics, which (for me) had mainly entailed the study of cognitive processes that were not directly observable. This had never bothered me; in fact, I had always maintained that I was only interested in the most fundamental kinds of psychological research, and that I was never going to 'sell my soul' and do applied research. However, more and more often I found myself looking at the world through a psychologist's eyes, wondering about the way in which many social processes went awry for a lack of psychological knowledge. Despite my earlier claims to the contrary, I had to admit that I was starting to get interested in the practical relevance of psychological research. Eventually I surprised my friends and family, and went to work in a department of Work and Organizational Psychology.

If I had expected the field of creativity research to be a natural mix of research and applications of research findings, I was up for a disappointment. In fact, particularly in the field of brainstorming, there often seems to be a large gap between what researchers study and what practitioners recommend. Research results apparently are not always taken to be directly relevant for professional practice, and anecdotal reports from professional practice are not usually taken to be very important for research. This was an important lesson for me. Studying creativity in controlled laboratory settings, and sometimes even finding nice (i.e., statistically significant) effects, it is easy to forget (or ignore) that even the most solid or elegant explanation may well meet with indifference (as illustrated by the first motto of this dissertation; fortunately, this particular quote did not concern my own research). On the other hand, creativity without the aid of well-developed guiding principles often leads to the situation that all kinds of wild and wonderful ideas or plans are formed and then abandoned (as illustrated by the second motto). To a certain degree, this will probably always be the case. Nevertheless, I hope that the studies reported in this dissertation may do something to help bridge the gap between creativity research and practice.

Writing a doctoral dissertation, like all forms of creative behavior, is strongly dependent upon the support that the author receives from colleagues, supervisors, and friends. Four years of doctoral research inevitably contain their share of frustration and doubts, but on the whole it has been fun and I would not hesitate to do it again (although I might do some things differently).

I want to thank my supervisors Bernard, Wolfgang, and Carsten for inspiring and instructing me for four years. All three of you have shaped my thinking about research and psychology in many ways, and have helped me with countless suggestions. I consider myself fortunate to have studied with you. Bernard: I particularly thank you for always having or making time for me, my sudden flashes of inspiration, and my SPSS printouts.

Thanks further go to all of my former colleagues in the Department of Work and Organizational Psychology at the University of Amsterdam for being such an inspiring, funny, and friendly group of people, and for putting up with my compulsive expressions of creativity. You have made me feel at home and have taught me a lot. Bianca: thank you for being prosocial and supportive. Irene: thank you for being a great roommate, even though I sometimes drove you crazy with my humming and foot-tapping.

Benjamin de Boer, Anna Borleffs, Mariëlle Chiang, Marieke Holterman, Rosa Noordam, Charlotte Pasteuning and Katherine Stroebe have helped with the collection and coding of the data reported in this dissertation. I thank you all for your patience and your conscientiousness.

Many of my friends and family members have contributed to this dissertation, if only by listening patiently to my enthusiastic stories – or complaints. I especially want to thank Ruth Zeifert for asking impossible questions, and for endless e-mail conversations; Maarten Muis for years of discussions, friendly disagreements and fits of laughter; Arne Moll for always being there, and for always understanding exactly what I mean, even before I have said it; and my father, Eric senior, for trusting in me and always giving me the freedom to do what I thought I should do. Most importantly of all: thank you, Esmee, for being my most critical reviewer and my strongest supporter. With this dissertation, as with all things, your comments and encouragement helped me to do things better – and your love kept me motivated to do them at all.

# Contents

CHAPTER 1

# Introduction: From Quantity to Quality

## 1.1 Introduction

Some concepts in psychology enjoy the general reputation of being even less scientific than the rest. Although most people would not deny that these concepts denote real phenomena, they are popularly supposed to be outside the realm of systematic study, are expected to fly in the face of cold scientific scrutiny, or are considered impossible to re-create in controlled laboratory settings. One of these is the concept of creativity. While creative behavior shapes almost all aspects of our society, including science, this type of behavior itself is often thought to be too elusive, too complex, or too mysterious for scientific explanation. Nevertheless, in this dissertation I will describe a number of scientific studies, report a fair amount of empirical data, and suggest some theory regarding creative behavior. Specifically, the central issue of these studies is the question whether *quantity breeds quality*: If we are trying to be creative and to come up with original and useful solutions to a problem, does it help to first think of as many ideas as we can? Does the generation of a large number of ideas make it more likely that we will eventually end up with a good idea?

The experiments reported in this dissertation are by no means the first scientific studies of creativity. In fact, a thriving line of psychological research argues strongly against the popular view that creativity is inherently unfit for scientific study. The starting point of systematic psychological research on creativity is conventionally marked by Guilford's (1950) Presidential Address to the American Psychological Association. In line with the above paragraph, Guilford stated: "I discuss the subject of creativity with considerable hesitation, for it represents an area in which psychologists generally, whether they be angels or not, have feared to tread" (p. 444). Guilford called for more research on creativity, and proposed that contemporary psychology already possessed the tools and concepts needed for such an endeavor. However, as a search of the PsycInfo database shows, of more than 15.000 records containing the word *creativity*, approximately two-thirds date from after 1980 (also see Feist & Runco, 1993). Thus, it seems to have taken psychologists some time to respond to Guilford's challenge.

This lack of attention from psychologists prior to 1980 does not mean that nobody was studying creativity at all. Shortly after Guilford's (1950) address, advertising agent Alex Osborn (1953; 1957; 1963[1]) published a book called "Applied imagination: Principles and procedures of creative problem-solving." This book contained many practical recommendations and procedures to increase creativity in companies and organizations, supported by the author's theory of creative behavior. One of these procedures remains well-known and highly popular today, under the name *brainstorming*. Many companies and organizations use brainstorming in order to come to innovations and creative solutions.

Brainstorming is an idea generation tool: People are supposed to come up with more creative ideas and solutions in a brainstorming session than by simply trying to solve a problem or to come up with some good ideas. According to Osborn, people have a natural tendency to be selective in the expression of their ideas. Thus, under normal circumstances, people always generate fewer ideas than they *could* have generated. Brainstorming is designed to remove this selective tendency and to maximize one's creative productivity. As the number of generated ideas or solutions increases, it becomes more likely that some of the best possible solutions are among those ideas. If, for example, only one in fifty ideas is an excellent idea, generating a hundred ideas yields a better chance of generating at least one excellent idea than generating twenty ideas.

However, the creative process does not end with idea generation. In fact, when brainstorming is used as an innovation tool, idea generation is only the beginning of the process (e.g., West, 2002). According to Osborn, generating more creative ideas will increase the chance that the best idea or ideas will actually be selected and implemented. It is this latter assumption that is central to this dissertation. Does generating a large number of ideas help the creative process, and if so, how? In what way does the quality of the ideas that people generate determine the quality of the ideas that they select as the "best" ones? Are people good at selecting creative ideas? However, I should first briefly address the question what brainstorming exactly is, and what is known about it.

[1] The 1953 edition is the first edition of the book (see reference list). I will mostly refer to the third, revised (1963) edition of Osborn's book, unless I am specifically discussing elements from an earlier edition.

## 1.2 Brainstorming

### *1.2.1 What Is Brainstorming?*

*Brainstorming Principles and Rules.* Although almost everybody is familiar with the term *brainstorming*, the word is used in many different ways. The term as Osborn (1963) used it (and as I will use it) denotes a specific procedure for idea generation, with specific rules. In a brainstorming session, the goal is to come up with as many creative ideas as possible. The main premise of the brainstorming procedure is that *quantity breeds quality*: The more ideas are generated, the more good ideas will be found among them. This relationship is assumed to be very strong, so that generating more ideas *always* means generating more good ideas. Thus, boosting productivity (i.e., the number of generated ideas) is the essence of brainstorming. "Always we should keep asking our imagination: 'What *else*?' and again, "What *else*?"' (Osborn, 1963, p. 130; italics in original). In order to reach this goal, participants in a brainstorming session should practice *deferment of judgment*: They should be free to express every idea they can think of, without worrying about possible criticism or evaluation by other participants (or themselves). If people refrain from expressing all their ideas, some potentially excellent ideas could be missed.

These two principles (quantity breeds quality, and deferment of judgment) are translated into four specific brainstorming rules: (1) *quantity is wanted*: express each and every idea that you can think of, (2) *criticism is ruled out*: do not criticize, judge, or evaluate ideas expressed by other participants or yourself (thus, remarks such as "that is impossible, it's never going to work" are strictly forbidden), (3) *freewheeling is welcomed*: do not worry about expressing strange, unconventional or bizarre ideas, but instead actively try to be as original and innovative as possible, and (4) *combine and improve*: feel free to adapt ideas that have been expressed earlier and to combine different ideas to come up with more innovative solutions.

Thus, the brainstorming rules are intended to remove inhibitions and to create an atmosphere of freedom and support, without the fear of being criticized or ridiculed. This atmosphere is supposed to be conducive to the production of ideas, and the production of ideas is supposed to lead to the production of *good* ideas.

*Brainstorming As Group Activity.* The word *brainstorming* in popular speech almost automatically denotes a group activity. The Oxford Advanced Learner's Dictionary (2000), for example, defines brainstorming as "a way of making a group of people all think about something at the same time, often in order to solve a problem or to create good ideas." Similarly, the Merriam-Webster Online Dictionary defines it as "a group problem-solving technique that involves the spontaneous contribution of ideas from all members of the group." Nevertheless, brainstorming can be done individually: Nothing in the four brainstorming rules prevents a person from conducting a brainstorming session by himself or herself.

One reason why brainstorming is often considered to be a group activity, is that people are thought to be more productive when brainstorming together. In the second edition of his book, Osborn (1957) included the prediction that people would be able to generate twice as many ideas when working in a group than when working alone. This prediction is intuitively appealing. Most of us have, at one time or another, come up with a new idea because of something somebody else said or wrote. Moreover, as popular wisdom has it, "two heads know more than one": Somebody else may know something we don't, and we may know something the other doesn't. It therefore seems sensible to generate ideas together, and most people believe that group brainstorming is more effective than individual brainstorming (Paulus, Dzindolet, Poletes, & Camacho, 1993; Stroebe, Diehl, & Abakoumkin, 1992).

### *1.2.2 Does Brainstorming Work?*

Research has shown that brainstorming is an effective idea generation procedure, and that the two basic principles of brainstorming indeed enhance the quantity and quality of generated ideas.

*Brainstorming and Idea Quantity.* In brainstorming, quantity (or productivity) is simply defined as the number of unique ideas that an individual or a group generates. Thus, after a brainstorming session, all the ideas generated by an individual or group are counted, with identical ideas only being counted once. Researchers have found in several studies that productivity goals (i.e., instructions to generate as many ideas as possible) significantly enhance productivity (e.g., Christensen, Guilford, & Wilson, 1957; Meadow, Parnes, & Reese, 1959; Shalley, 1991); thus, when people are instructed to generate as many ideas as possible, they do in fact generate more ideas than when they receive no such instructions.

*Brainstorming and Idea Quality.* Most researchers agree that, in the context of creativity, quality is a combination of *originality* (the degree to which a product or an idea is novel, unusual or innovative) and some measure of *appropriateness* (the degree to which a product or an idea is relevant to the topic, is expected to actually solve a problem, or is thought to be practically feasible). In fact, originality and appropriateness are usually taken to be the defining characteristics of creative behavior and creative products (e.g., Amabile, 1996; Diehl & Stroebe, 1987; Sternberg & Lubart, 1999). Thus, in brainstorming research, idea quality is generally measured by having one or more judges rate the originality and the feasibility (the degree to which the ideas can plausibly be realized) of the generated ideas. An idea must be original in order to be called creative, and it must be feasible in order to be more than just unusual or bizarre.

In line with Osborn's (1963) reasoning, research has confirmed that quantity indeed breeds quality; specifically, a higher productivity is associated with a higher availability of high-quality ideas or solutions (e.g., Diehl & Stroebe, 1987; Parnes & Meadow, 1959). A 'good' or high-quality idea is usually defined as an idea that scores above an arbitrary cutoff point on both originality and feasibility. Brainstorming studies typically find high positive correlations between the total number of generated ideas and the number of good ideas contained therein (e.g., $r$ = .82 in Diehl & Stroebe, 1987; $r$ = .69 in Parnes & Meadow, 1959).

*Deferment of Judgment.* Originality, one of the defining characteristics of creativity, essentially means deviation from a norm or convention. However, people have a strong tendency toward conformity (e.g., Asch, 1955) and deviants in a group are generally not very well-liked (e.g., Schacter, 1951). In other words: it is risky to be original. Signs that originality is not appreciated, such as criticism, are therefore likely to inhibit creative behavior. Indeed, people are generally more creative when they feel free of criticism or evaluation (e.g., Amabile, 1979; Hennessey, 1989), and idea generation is no exception to this rule. For example, Bartis, Szymanski, and Harkins (1988) found that deferment of judgment increased the average quality of the ideas generated by their participants, and Dentler and Mackler (1964) found that participants performed better on an "unusual uses" task (in which participants are instructed to generate as many unconventional uses as possible for a familiar object) when they felt safe and relaxed than when they had to work under challenging and competitive circumstances. Furthermore, Camacho and Paulus (1995) found a negative relationship between group members' degree of social anxiousness and brainstorming performance, and Diehl and Stroebe (1987) found that individual brainstormers were less productive when they believed that their performance would be judged.

Thus, both brainstorming principles are supported by empirical data: People generate more ideas, and more high-quality ideas, when they are trying to generate as many ideas as possible, and people are more creative when they do not feel constrained by evaluation.

*Productivity and Productivity Loss in Interactive Groups.* In contrast, Osborn's (1957) prediction that group brainstorming would be much more effective than individual brainstorming has received less empirical support. In fact, starting with a paper published by Taylor, Berry, and Block (1958), a long line of studies has shown that people are not more, but *less* productive when they are brainstorming in a group (see Diehl & Stroebe, 1987, Lamm & Trommsdorff, 1973, and Mullen, Johnson, & Salas, 1991, for overviews). In this line of research, a comparison is usually made between groups of people brainstorming together (*interactive* groups) and the same number of people brainstorming individually, whose unique ideas are added together after the brainstorming session (*nominal* groups). It is consistently found that interactive groups generate fewer ideas, and fewer high-quality ideas, than nominal groups. This effect is quite strong: Nominal groups can easily generate twice as many ideas as interactive groups. Furthermore, the productivity loss in interactive groups increases with group size, so that an interactive group of six people suffers a larger productivity loss than an interactive group of only two people. Thus, contrary to Osborn's prediction, group brainstorming (as compared to individual brainstorming) does not make people more productive. In fact, the opposite is the case.

*Causes of Productivity Loss in Groups.* Many of the existing brainstorming studies have addressed this productivity loss in interactive groups, studying possible causes and remedies. These studies have revealed several factors that can contribute to productivity loss, such as *social inhibition*: Despite the fact that the brainstorming instructions explicitly forbid criticism, participants in a group brainstorming session may nevertheless fear other participants' judgment or disapproval, and therefore fail to express all their ideas (e.g., Camacho & Paulus, 1995; Collaros & Anderson, 1969). Alternatively, instead of worrying about possible evaluation, group members may feel that their individual contribution to the group output cannot be identified, and this may give rise to *free riding* or *social loafing*: a tendency to dispense less effort on idea generation and to let the other people in their group do most of the work (e.g., Diehl & Stroebe, 1987).

Although these and other factors can all inhibit productivity in brainstorming groups, a number of studies (Diehl & Stroebe, 1987, 1991; Nijstad, Stroebe, & Lodewijkx, 2003) suggest that another factor, *production blocking*, is the main cause of productivity loss in interactive groups. In brainstorming groups, only one person can speak at a time, so that when one group member is expressing an idea, other group members have to listen to that idea. During that time, they cannot generate new ideas. Even worse, ideas that they had already generated (but not yet expressed) are likely to be forgotten. Thus, group interaction directly interferes with the cognitive processes underlying idea generation. When one is brainstorming alone, no such blocking and interference is possible.

### *1.2.3 Is Brainstorming Useful?*

The effectiveness of the brainstorming rules has been confirmed empirically, but that is not to say that the effectiveness of brainstorming as an innovation tool has been established beyond doubt. There is more to creativity and innovation than idea generation.

*The Context of Brainstorming.* As explained earlier, the goal of a brainstorming session is to generate as many creative ideas as possible; thus, the emphasis is on productivity. "It is almost axiomatic that quantity breeds quality in ideation. Logic and mathematics are on the side of the truth that the more ideas we produce, the more likely we are to think up some that are good" (Osborn, 1963, p. 131). As explained above, research has corroborated this hypothesis. Since interactive brainstorming groups display a considerable productivity loss, their contribution to the innovation process is thought to be limited, and brainstorming researchers usually recommend against the use of interactive brainstorming groups. However, brainstorming does not happen in a vacuum, and producing large numbers of ideas is never the ultimate goal of a brainstorming session. Instead, what brainstormers are after is a limited number of good ideas to select for further development and, eventually, implementation (Nijstad & De Dreu, 2002). Thus, when brainstorming is regarded as an innovation tool, a focus on productivity is inadequate, and so are recommendations exclusively based on idea quantity (see Sutton & Hargadon, 1996, for a similar view).

*The Stages of Creative Problem-Solving.* Although the brainstorming procedure is sometimes seen, and certainly often studied, as a creativity-enhancement technique in itself, this view is not quite correct. Osborn (1963) considered brainstorming to be only part of the creative problem-solving process. In his view, this process comprised three stages: (1) *fact finding* (defining a problem, gathering and analyzing relevant data), (2) *idea finding* (generating, selecting and modifying ideas), and (3) *solution finding* (evaluating, testing and adopting solutions). Osborn's views were later elaborated into the six-stage Osborn-Parnes Creative Problem Solving model (CPS; e.g., Parnes, 1992), which remains highly influential. The six stages in this model are: (1) *goal finding* (definition of the goal or objective that one wants to meet), (2) *fact finding* (gathering goal-relevant data), (3) *problem finding* (formulating the problem definition), (4) *idea finding* (generating creative ideas), (5) *solution finding* (formulating or revising the criteria by which the ideas or solutions will be judged), and (6) *acceptance finding* (gathering support for the selected solution and devising a plan of action for development and implementation).

Such multi-stage views of creativity are quite common. Many theories or models of creativity divide the creative process into distinct stages, especially with regard to applied creativity or innovation (e.g., Amabile, 1996; Baer, 2003; Basadur, 1995; Perkins, 1981; Révész, 1952; West, 2002). It seems somewhat strange, then, that brainstorming research has largely neglected the context in which a brainstorming session takes place. To a certain degree, this can probably be explained by the initially surprising findings with regard to productivity loss in groups, and the effort that researchers have invested in uncovering the associated causes and possible remedies. However, after almost fifty years of brainstorming research, a more inclusive research strategy is desirable, both from a theoretical and an applied perspective.

### *1.2.4 From Quantity to Quality*

If one approaches the question of brainstorming effectiveness with Osborn's (1963) three stages of creative problem solving in mind, it is clear that brainstorming is effective to the extent that it leads to the *selection* and subsequent *implementation* of high-quality solutions. The degree to which this is the case depends not only on the availability of high-quality ideas, but also on the selection process itself. It may be well-established that quantity (number of available ideas) and quality (number of available high-quality ideas) are strongly correlated in idea generation, but this does not mean that quantity is as strongly related to quality in the stage *after* idea generation. In other words: although generating more ideas increases the chance that the best idea will be among these ideas, it remains an empirical question whether this also increases the chance that the best idea will be recognized as such and selected for further development. One aim of this dissertation is to provide an answer to this question.

Previous brainstorming studies have focused on the productivity difference between interactive and nominal groups, and the conclusion that is generally drawn from these studies is that interactive groups should not be used. However, such a recommendation is premature if we do not know whether the lower productivity of interactive groups also causes them to select ideas of lower quality. Because this question follows directly from the existing line of brainstorming research and has important implications for practice, this is the first question that this dissertation will address.

Having established whether the higher productivity of nominal groups also leads to the selection of better ideas, I will turn to the factors underlying creative idea selection after brainstorming. The first of these factors is the way in which quantity leads to quality. Especially when seen in the context of idea selection and innovation, the relationship between quantity and quality clearly is an important one. However, it is not well understood. The strong correlation usually found between productivity and number of good ideas implies a linear relationship: As total productivity increases, the number of available high-quality ideas increases proportionally, while the average quality of the ideas does not change. A parsimonious interpretation of this relationship is that idea generation is a chance or stochastic process, in which every single idea that is generated has an equal chance of being a good, creative idea. If so, the only way to increase the availability of good ideas is to increase total productivity. Alternatively, if idea generation is *not* a chance process, and if the average quality of the generated ideas is not really independent of productivity, this could have important implications for the brainstorming procedure and for creativity research in general. Thus, it is important to know whether the relationship between quantity and quality is governed by mere chance. This is the second question that this dissertation will address.

The other essential factor underlying creative idea selection is, of course, the effectiveness of the selection process itself. Whether or not the relationship between quantity and quality is best described as a chance process, the available research results suggest that a higher productivity is generally associated with a higher availability of both high- and low-quality ideas. Consequently, brainstormers must be able to recognize and select the best, or better-than-average ideas, if brainstorming is to be effective in a non-trivial sense. However, because little research on this issue is available, we do not know how far the motto "quantity breeds quality" really extends. Thus, the third and final issue addressed in this dissertation is the question whether people are actually able to recognize and select their most creative ideas, how they go about making such a selection, and how the selection of creative ideas might be improved. I will specifically focus on the degree to which creative idea selection benefits from a thorough consideration of the available ideas and the use of explicit selection criteria.

### 1.3 Summary and Overview

Brainstorming, an idea generation technique developed by Osborn (1953), rests on the premise that quantity breeds quality: Generating more ideas means generating more good ideas, and a higher availability of good ideas in turn should increase the chance that a good idea will be selected for further development. However, the latter part of this argument has hardly been studied. Thus, although we know a good deal about group and individual brainstorming and the processes involved (such as the processes leading to production loss in brainstorming groups), we do not know very much about the degree to which the generation of creative ideas in a brainstorming session actually contributes to the selection of high-quality solutions.

Although the brainstorming procedure presupposes that the most productive brainstormers will also come up with the best solutions, we do not know whether the most productive groups or individuals will also prove to select the best ideas. This question regarding selection outcome, and particularly the question whether nominal groups' higher productivity will also cause them to select better ideas than interactive groups, will be the central topic of the studies reported in Chapter 2. Next, Chapter 3 will address the quality of ideas that are generated in a brainstorming session, with emphasis on the question whether or not the relationship between quantity and quality is governed by chance. Finally, although the availability of creative ideas is a necessary condition for creative idea selection, the nature of this selection process and the factors influencing it are unknown. The studies reported in Chapter 4 aim to shed light on the process of idea selection. Together, the studies reported in this dissertation should lead to a more complete understanding of brainstorming effectiveness, as well as to practical recommendations regarding applied creativity.

CHAPTER 2

# Idea Selection in Interactive and Nominal Groups [1]

## 2.1 Introduction

Although brainstorming studies have yielded important insights into the nature of idea generation and into the possible stumbling blocks that brainstorming practitioners may encounter, the literature is remarkably silent when it comes to the selection of creative ideas. Thus, while most brainstorming researchers would strongly recommend against the use of face-to-face brainstorming groups, this recommendation is based only on data regarding productivity: Interactive groups generate fewer ideas than people working individually, whose ideas are then added together (nominal groups). This is a rather limited approach. If productivity was the only goal of a brainstorming session, group brainstorming indeed should probably be avoided or minimized. However, in real life, the creative process does not end with the generation of creative ideas.

As several authors have noted, creativity is not a single act, but consists of multiple stages (e.g., Baer, 2003; Basadur, 1995; Osborn, 1963). As explained in the previous chapter, Osborn (1963), in his description of the brainstorming procedure, made it clear that idea generation is only part of the creative process. This is particularly relevant if brainstorming is used as an innovation tool: Although the availability of creative ideas is a necessary condition for innovation, it is not sufficient (e.g., Nijstad & De Dreu, 2002; West, 2002). For creativity to become innovation, creative ideas must be implemented, and implementation of creative ideas can only take place after they have been selected from the total pool of generated ideas. Effective selection of ideas after a brainstorming session is therefore essential. However, little research has been done on the issue of creative idea selection.

Given the neglect of idea selection as a research topic, I agree with Sutton and Hargadon (1996) that "it is premature to conclude that face-to-face brainstorming groups are ineffective" (p. 688). Before passing such a harsh judgment, it should first be assessed whether the higher productivity of nominal groups also leads to the selection of better ideas. If nominal groups generate a larger number of creative ideas, but do not manage to

[1] This chapter is based on Rietzschel, Nijstad, and Stroebe (in press).

make good use of these ideas and end up selecting the same ideas (or ideas of similar quality) as interactive groups, there seems to be little necessity to refrain from group brainstorming. Alternatively, if interactive groups not only generate fewer ideas than nominal groups, but also select ideas of lower quality, the recommendation against the use of interactive brainstorming groups is supported from a new perspective.

In this chapter, I will go into the issue of idea selection. Specifically, I will address three questions: (1) the question whether the higher productivity of nominal groups also leads to their selecting better ideas, (2) the question how good interactive and nominal groups are at recognizing and selecting creative ideas, and (3) the question whether idea generation and selection should be strictly separated, as is traditionally recommended in brainstorming.

### *2.1.1 Idea Selection in Interactive and Nominal Groups*

The outcome of idea selection is necessarily dependent on two factors: the *availability* of good ideas, and the *effectiveness* of the selection process. Clearly, if there are no good ideas available to choose from, it will by definition be impossible to select good ideas. Since the goal of a brainstorming session is to boost idea generation and thereby boost the production of good ideas, a successful brainstorming session should lead to the availability of many good ideas, facilitating the selection of good ideas. Further, regardless of the actual quality of the available ideas, idea selection can be more or less effective, depending on the degree to which people selected the best ideas possible. For example, the selection of a moderately good idea is a very effective selection if all other ideas are of very low quality, but a very ineffective selection if all other ideas are of very high quality. Thus, although the quality of selected ideas is necessarily *constrained* by the quality of the available ideas, it is not directly *determined* by it. The ideas that get selected should be the best available ideas, or at least be among the best ideas available.

*Availability of Good Ideas.* What exactly is a 'good' idea? In brainstorming studies, idea quality is usually defined as a combination of originality and feasibility (e.g., Diehl & Stroebe, 1987). This fits with the approach generally taken by creativity researchers. For example, Sternberg and Lubart (1999) define creativity as "the ability to produce work that is both novel (i.e., original, unexpected) and appropriate (i.e., useful, adaptive concerning task constraints)" (p. 3). Similarly, Amabile (1996) defines a creative response to a task as one that is "both a novel and appropriate, useful, correct or valuable response to the task at hand" (p. 35). While originality is often considered to be the hallmark of creative behavior, some form of appropriateness is essential to distinguish truly creative behavior from behavior that is merely unusual or erratic. Likewise, although ideas must be original in order to be called creative, they will not be implemented if they are not feasible. Hence, the usual definition of a good idea is an idea that is both highly original (or unusual), and highly feasible (or useful).

With regard to the availability of good ideas, the superiority of nominal groups over interactive groups is clear. Several brainstorming studies have taken idea quality into account, and have shown that the number of good ideas increases more or less linearly with the total number of generated ideas. For example, Parnes and Meadow (1959) found a strong positive correlation between the total number of generated ideas and the number of good ideas (i.e., ideas that were judged by a rater to be both highly original and highly useful). Similarly, Diehl and Stroebe (1987) found that the higher productivity of nominal groups as compared to interactive groups was accompanied by the production of more good ideas. Thus, in line with the brainstorming assumptions, enhancing productivity enhances the availability of good ideas. Nominal groups are more productive than interactive groups, and hence produce more high-quality ideas. However, it is not immediately obvious that this will automatically increase the quality of ideas selected by nominal groups.

*Selection Effectiveness.* Group interaction could have positive or negative effects on the effectiveness of the selection process. On the one hand, various sorts of process loss can lead to suboptimal idea selection in interactive groups (also see Chapter 1). For example, decision making groups often fail to discuss all relevant information and to consider all available alternatives (see Stasser, 1999). Of course, effective selection largely rests on a thorough consideration of available options. If interactive groups do not consider the options at their disposal, they cannot be expected to compensate for their lower productivity. It would be reasonable, therefore, to expect that the quality of selected ideas will be higher for nominal groups than for interactive groups.

On the other hand, several studies have demonstrated that groups can outperform individuals, particularly on intellective tasks (task with demonstratively correct answers). For example, Laughlin and colleagues (Laughlin, Bonner, & Altermatt, 1998; Laughlin & Shippy, 1983; Laughlin, VanderStoep, & Hollingshead, 1991) found that groups performed better than individuals on a task in which participants had to decide between hypotheses to account for certain patterns. Although idea selection is not an intellective, but a judgmental task, group discussion clearly can improve the selection process. Thus, one could also predict that interactive groups will overcome their lower productivity, and will select ideas that are at least as good as those selected by nominal groups.

*Previous Research.* Recently, some studies have addressed the selection of creative ideas after a brainstorming session, and the effects of nominal versus interactive brainstorming on idea selection.

Putman and Paulus (2003) conducted a study in which participants were instructed to generate and select ideas. Half of the participants generated ideas in three-person interactive groups, and half of the participants generated ideas individually, as members of nominal groups. After the brainstorming task, members of interactive groups selected the five best ideas from their group's production, while members of nominal groups were placed together and selected the best five ideas from their pooled production. All ideas received two originality scores (one rated on a 5-point scale by a trained rater, and one based on the frequency of occurrence of the idea in the overall idea pool) and a feasibility score (as rated on a 5-point scale by a trained rater). Putman and Paulus found that, as usual, interactive groups generated fewer ideas than nominal groups. They also found that the ideas generated by nominal groups were, on average, more original than those generated by interactive groups. Furthermore, the average originality of the five selected ideas was higher for groups who had generated ideas as nominal groups than for groups who had generated ideas as interactive groups. Thus, those groups who had generated ideas individually did manage to reap the benefits of their higher productivity, and selected ideas that were higher in originality. For feasibility, no such effects were found.

As a measure of selection effectiveness, Putman and Paulus (2003) computed the average originality (resp. feasibility) of the five ideas selected by each group, and divided this by the average originality (resp. feasibility) of the five most original ideas generated by that group (i.e., the five most original (resp. feasible) ideas that the group could have selected). Theoretically, these ratios could take values between .2 and 1. According to this measure, nominal and interactive groups did not differ in the effectiveness of their selection. It is hard to interpret the size of these ratios, but the fact that the reported values lie between .60 and .70 suggests that selection effectiveness was not optimal.

Faure (2004) conducted a similar study, but with different results. Participants in this study first generated ideas, either as an interactive group or individually as a nominal group. Next, every participant made an individual preselection of his or her five favorite ideas from the group's total production (i.e., all the ideas generated by that participant's nominal or interactive group). Finally, the group members together selected the three best ideas from the preselected ideas. All ideas received dichotomous scores for originality (ideas that were mentioned by less than 10% of all groups were counted as original, and all other ideas were counted as unoriginal), and were rated by four judges on 5-point scales for practicality and effectiveness. In accordance with Putman and Paulus (2003), Faure found that nominal groups generated a larger number of ideas, and a larger number of original ideas, than interactive groups. However, she found no difference between interactive and nominal groups with regard to the quality of selected ideas. In other words, the higher productivity of nominal groups, and the resulting higher availability of original ideas, did not lead to the selection of better ideas. No measure of selection effectiveness was reported in this paper.

The picture that emerges from these results is that, although productivity does enhance the potential for creative idea selection, a high productivity clearly does not automatically lead to the selection of creative ideas. However, Faure (2004) reports no measure of selection effectiveness, and the measure reported by Putman and Paulus (2003) is difficult to interpret. Thus, we do not know whether idea selection in interactive and nominal groups is effective or not. This is an important question, because the availability of high-quality is useful only if people are able to recognize and select their best ideas. If selection effectiveness is very low, this could limit the practical use of brainstorming as an innovation tool.

Furthermore, while both of these studies aimed to make a comparison between interactive and nominal groups, all groups in these studies selected ideas as interactive groups. Thus, as yet, no actual comparison has been made between interactive and nominal groups on idea selection. It therefore remains unknown whether nominal groups select better ideas than interactive groups if they also select ideas as a nominal group, and whether interactive or nominal groups are more effective in the selection process.

### *2.1.2 Combining Idea Generation and Idea Selection*

Even though idea generation and idea selection are both essential parts of the innovation process, combining them effectively is considered an onerous task. Brainstorming conventions dictate that idea selection should be separated from idea generation as strictly as possible. An important “active ingredient” (Smith, 1998) in the brainstorming procedure is the deferment of judgment: People are thought to generate more ideas, or more creative ideas, when they feel free of evaluation and criticism (e.g., Amabile, 1979; Hennessey, 1989), so separation of idea generation and selection is essential. The brainstorming literature provides some indirect support for this assumption (e.g., Camacho & Paulus, 1995; Diehl & Stroebe, 1987), but the effects of a separation between idea generation and idea selection have not been directly studied.

In line with previous findings, it is to be expected that removing this strict task separation will make the evaluative aspect inherent in idea selection more salient during idea generation, and that this will inhibit productivity. This effect should be especially strong in interactive groups, for two reasons: Firstly, only in interactive groups can members be evaluated by others; thus, one should expect any manipulation that increases evaluation apprehension (the fear that one's ideas will receive a negative evaluation; also see Chapter 1) to have a stronger effect for interactive than for nominal groups. Secondly, interactive groups suffer from production blocking: Members of interactive groups have to take turns expressing their ideas, and this interferes with group members' idea generation and expression (Diehl & Stroebe, 1991). Removing the task separation should exacerbate this production blocking, because group members will devote more time to explaining their ideas, which means that other group members will have to wait longer before they can express their ideas.

The presence or absence of a strict task separation could influence idea selection in two ways. If the separation of idea generation and selection enhances productivity, more ideas are available for selection; this in turn could lead to the selection of better ideas as compared to the situation in which generation and selection are not strictly separated. However, because removing the task separation should cause brainstormers to devote more time to idea evaluation, even while still generating ideas, it could also improve the selection process. Therefore, it is possible that even though productivity will be negatively affected by removing the traditional task separation, the quality of the selected ideas will not be negatively affected.

## 2.2 Study 2.1: Idea Selection in Interactive and Nominal Groups

### *2.2.1 Introduction*

The primary aim of this study was to make a direct comparison between interactive and nominal groups on a task that involved both idea generation and idea selection. Participants in this task first generated creative ideas in a brainstorming task, and then selected the best ideas from their own production.

I also manipulated task separation, by presenting idea generation and idea selection either as two separate tasks, or as one task consisting of two activities (idea generation and selection).

I expected nominal groups to generate more ideas, and more high-quality ideas, than interactive groups, especially in the absence of a strict separation between idea generation and selection. However, given the lack of research on idea selection by interactive and nominal groups, I had no specific expectations with regard to the quality of the selected ideas.

*2.2.2 Method*

*Participants*

In total, 138 students of the University of Amsterdam participated in the experiment (102 females and 36 males, mean age = 21.4 years). Participants received course credit or 10 Euros (about 13 US Dollars). All participants were assigned to three-person groups that were homogeneous with regard to gender, yielding 12 male groups and 34 female groups. Four groups were excluded from the analyses, because not all of the group members followed task instructions.

*Independent Variables*

The study used a 2 (Type of Group: interactive versus nominal) x 2 (Task Separation: two tasks versus one task) factorial design. Participants were assigned randomly to conditions.

*Procedure*

First, all participants read a general introduction in which it was explained that they would participate in a study regarding creativity and innovation, and that they would generate and select ideas. During idea generation and selection, members of interactive groups worked together in one room; members of nominal groups worked alone in separate rooms. All sessions were videotaped, in order to measure brainstorming times.[2]

The experimental task consisted of two parts: idea generation and idea selection. In the first part, participants were asked to generate ideas about ways to improve education at the Department of Psychology. The four brainstorming rules were provided: (1) generate as many ideas as possible, (2) freewheeling is encouraged, (3) criticism and evaluation are not allowed, and (4) combine and improve ideas. Participants wrote down their ideas on A4-sized sheets of paper, which were divided into sections to facilitate the counting of ideas. In interactive groups, all ideas were verbalized and then written down by one of the group members. In the idea selection part of the task, participants selected and rank ordered the four best ideas; these were to be written down on a designated A4-sized sheet of colored paper.

[2] In the two-task condition, brainstorming time was defined as the period between the moment the experimenter left the room and the moment the group or individual stopped brainstorming. In the one-task condition, this boundary could not be drawn so easily. I therefore defined brainstorming time as the period between the moment the experimenter left the room and the moment when the group or individual visibly switched to idea selection by reaching for the selection sheet and leafing through the ideas.

Participants in the two-task condition first performed the idea generation task; after 30 minutes, or when participants indicated that they had finished generating ideas, the experimenter handed out the instructions and materials for the selection task. For this task, a maximum of 30 minutes was available as well.

In the one-task condition, participants received the instructions for idea generation and idea selection simultaneously. The content of these instructions was the same as in the two-task condition (including the four brainstorming rules), but they were presented together, as one set of instructions. Thus, the difference between the two conditions was that idea generation and selection were either framed as two separate tasks, or as one task consisting of two activities. I did not give participants in the one-task condition the instruction to stop brainstorming after 30 minutes, because this would interfere with my manipulation. Consequently, whereas participants in the two-task condition had 30 minutes for idea generation and 30 minutes for idea selection, participants in the one-task condition simply had 60 minutes for the whole task. Note that participants in the one-task condition did not receive any instructions with regard to task separation; thus, they were free to switch between idea generation and selection, or to separate the tasks.

In all conditions, idea selection was followed by a questionnaire that participants completed individually. Participants were then thanked, debriefed and paid.

*Measures and Dependent Variables*

*Productivity.* As usual, productivity was defined as the number of non-redundant ideas per group. For nominal groups, the ideas produced by three individuals were added together. Redundant ideas, defined as identical ideas within one group's production, were counted only once.

*Idea quality.* A trained rater who was blind to conditions coded all unique ideas for originality and feasibility (see Appendix 1). For both dimensions, a 5-point scale was used (1 = *not original* [*feasible*], 5 = *highly original* [*feasible*]). Examples of a highly original and a highly unoriginal idea, respectively, are: "Use hypnosis to increase students' concentration," and "Teach courses in smaller groups." Examples of a feasible and an unfeasible idea, respectively, are: "Maintain a stricter policy against cell phones during exams," and "Make all course books available in digital form." A second rater scored 10% of the ideas to determine interrater reliability. With raters considered to be in agreement whenever ratings differed by no more than one point (Diehl & Stroebe, 1987), agreement existed in 97.7% of the cases for originality, and in 97.2% of the cases for feasibility. I also computed intraclass correlations, using a two-way random model and consistency definition (McGraw & Wong, 1996; Shrout & Fleiss, 1979); these were .77 for originality and .73 for feasibility. These values are "good" to "excellent" according to criteria specified by Cicchetti and Sparrow (1981). Hence, I used the scores of the first rater in my analyses.

I computed the mean originality and the mean feasibility of the generated ideas and the selected ideas. Interactive groups selected the four best ideas as a group; in nominal groups, each group member selected his/her four best ideas as an individual, yielding twelve selected ideas per nominal group. Having participants use a rank ordering procedure in the selection task allowed me to compute quality scores for the *three favorite ideas* in each interactive and nominal group, that is, the three ideas that received the highest ranking within each group. For the interactive groups, I computed the mean originality and feasibility scores of the selected ideas with rankings 1, 2 and 3; for the nominal groups, I computed the mean originality and feasibility of the three selected ideas with ranking 1. I also computed alternative measures for the quality of the selected ideas, such as the mean originality and feasibility of *all* selected ideas, but these measures showed the same results in the analyses and hence are not reported here.

For each interactive and nominal group, I also computed the number of generated high-quality ideas, with a high-quality idea defined as an idea that received an originality score that was above the average originality score in this study and a feasibility score above the average feasibility score.

*Satisfaction.* In the questionnaire, I asked participants how satisfied they felt about their productivity, the quality of their ideas, the quality of the selection process and the quality of the selected ideas. Responses to all these items were given on a 9-point scale (1= *not at all satisfied*, 9 = *very satisfied*).

### *2.2.3 Results*

#### *Productivity*

Table 2.1 presents the means and standard deviations for productivity and the quality of generated and selected ideas. A 2 (Type of Group) x 2 (Task Separation) ANOVA on productivity yielded a main effect for type of group ($F(1, 38) = 58.53$, $p < .001$, partial $\eta^2 = .61$): Nominal groups generated more ideas ($M = 57.05$, $SD = 17.85$) than interactive groups ($M = 27.45$, $SD = 10.35$). A main effect of task separation ($F(1, 38) = 5.64$, $p = .023$, partial $\eta^2 = .13$) was qualified by a Type of Group x Task Separation interaction ($F(1, 38) = 9.66$, $p = .004$, partial $\eta^2 = .20$). Simple effects analysis showed that one-task nominal groups generated more ideas than two-task nominal groups ($F(1, 38) = 14.40$, $p = .001$); this difference was not significant for the interactive groups ($p > .5$).

To test whether the high productivity of one-task nominal groups was caused by the longer time available for brainstorming in the one-task condition (theoretically, one-task groups could spend 59 minutes brainstorming and 1 minute selecting ideas[3]), I divided the number of ideas generated in each group by the number of minutes spent brainstorming by that group (measured from the video recording of each session; for two groups, no video recordings were available, due to equipment failure). An ANOVA yielded a main effect for type of group ($F(1, 36) = 31.21$, $p < .001$, partial $\eta^2 = .46$): Nominal groups generated more ideas per minute ($M = 1.98$, $SD = 0.51$) than interactive groups ($M = 1.10$, $SD = 0.62$). The interaction Type of Group x Task Separation was significant ($F(1, 36) = 8.35$, $p = .006$, partial $\eta^2 = .19$); simple effects analysis showed that, as predicted, the one-task interactive groups generated fewer ideas per minute than two-task interactive groups ($F(1, 36) = 5.68$, $p = .023$). One-task nominal groups, however, generated slightly more ideas per minute than two-task nominal groups (although this effect was only marginally significant: $F(1, 36) = 2.96$, $p = .094$).

Table 2.1 - *Means and Standard Deviations for Idea Quantity and Quality in Study 2.1*

| | 1 Task | | 2 Tasks | |
|---|---|---|---|---|
| Measure | Nominal group | Interactive group | Nominal group | Interactive group |
| Number of ideas | 67.7 (13.38) | 25.9 (6.79) | 46.4 (15.54) | 28.75 (12.75) |
| Number of high-quality ideas | 15.80 (5.16) | 4.70 (3.65) | 10.40 (5.17) | 3.75 (2.63) |
| Number of ideas per minute[a] | 2.20 (0.45) | 0.78 (0.21) | 1.79 (0.50) | 1.34 (0.72) |
| Originality of generated ideas[b] | 1.61 (0.12) | 1.44 (0.22) | 1.51 (0.28) | 1.37 (0.25) |
| Feasibility of generated ideas[b] | 3.71 (0.21) | 3.98 (0.44) | 3.67 (0.19) | 3.90 (0.31) |
| Originality of selected ideas[b] | 1.47 (0.23) | 1.67 (0.59) | 1.53 (0.48) | 1.44 (0.29) |
| Feasibility of selected ideas[b] | 3.77 (0.39) | 3.90 (0.42) | 3.93 (0.44) | 3.67 (0.62) |

*Note.* $N = 42$ groups. Standard deviations are in parentheses. Originality and feasibility measures refer to *average* originality and feasibility.
[a]$N = 40$ groups.
[b]Maximum value = 5.

*Idea Quality*

I performed a MANOVA with a 2 (Quality: originality vs. feasibility) x 2 (Activity: production vs. selection) x 2 (Type of Group: interactive vs. nominal) x 2 (Task Separation: two tasks vs. one task) mixed design. Quality and activity were within-subjects factors; type of group and task separation were between-subjects factors. The analysis

[3] An ANOVA on the number of minutes spent brainstorming revealed a main effect for task separation ($F(1, 36) = 12.39$, $p = .001$, $\eta^2 = .26$): one-task groups spent more time in idea generation ($M = 31.59$, $SD = 7.25$) than two-task groups ($M = 24.32$, $SD = 5.48$). There were no other significant effects (*p*s > .1).

revealed a significant main effect of quality ($F(1, 38) = 950.09, p < .001$, partial $\eta^2 = .96$): The generated and selected ideas were higher in feasibility ($M = 3.81$, $SD = 0.37$) than in originality ($M = 1.49$, $SD = 0.30$).

Surprisingly, there was no effect of activity, indicating that the average quality of the selected ideas was not significantly different from the average quality of the generated ideas ($p > .50$). I also found a significant Quality x Activity x Type of Group interaction ($F(1, 38) = 8.79, p = .005$, partial $\eta^2 = .19$). In order to analyze the interaction, I performed separate analyses for originality and feasibility.

*Originality.* A 2 (Originality: originality of the generated ideas vs. originality of the selected ideas) x 2 (Type of Group) x 2 (Task Separation) mixed ANOVA revealed a marginal interaction of Originality x Type of Group ($F(1, 38) = 2.97, p = .093$, partial $\eta^2 = .07$). Further analysis showed that nominal groups generated more original ideas ($M = 1.56$, $SD = 0.21$) than interactive groups ($M = 1.40$, $SD = 0.23$; $F(1, 38) = 4.87, p = .033$, partial $\eta^2 = .11$), but that this difference was not present for the selection task ($p > .50$). The disappearance of this difference was due to the interactive groups selecting ideas that were slightly more original ($M = 1.55$, $SD = 0.45$) than their average production ($M = 1.40$, $SD = 0.23$), although this difference was only marginally significant ($F(1, 38) = 3.16$, $p = .083$). For nominal groups, no such difference was found.

*Feasibility.* A 2 (Feasibility: feasibility of the generated ideas vs. feasibility of the selected ideas) x 2 (Type of Group) x 2 (Task Separation) mixed ANOVA revealed a significant interaction of Feasibility x Type of Group ($F(1, 38) = 5.69, p = .022$, partial $\eta^2 = .13$). Interactive groups generated more feasible ideas ($M = 3.94$, $SD = 0.37$) than nominal groups ($M = 3.69$, $SD = 0.20$; $F(1, 38) = 6.93, p = .012$, partial $\eta^2 = .15$), but this difference disappeared in the selection task ($p > .5$). As with originality, this interaction effect was driven by the interactive groups; interactive groups selected ideas that were slightly (but not significantly) less feasible ($M = 3.77$, $SD = 0.54$) than their average production ($M = 3.94$, $SD = 0.37$; $F(1, 38) = 3.0, p = .091$). Again, no such difference was found for the nominal groups.

*High-Quality Ideas.* A 2 x 2 ANOVA with the number of high-quality ideas as dependent variable yielded a main effect of type of group ($F(1, 38) = 46.18, p < .001$, partial $\eta^2 = .55$), indicating that nominal groups generated more high-quality ideas ($M = 13.10$, $SD = 5.74$) than interactive groups ($M = 4.18$, $SD = 3.09$), and a main effect of task separation ($F(1, 38) = 5.91, p = .02$, partial $\eta^2 = .14$), indicating that one-task groups generated more high-quality ideas ($M = 10.25$, $SD = 7.17$) than two-task groups ($M = 6.77$, $SD = 5.15$). This latter effect was qualified by a marginally significant interaction of type of group and task separation ($F(1, 38) = 2.902, p = .09$, partial $\eta^2 = .07$). Simple effects analysis showed that this effect of task separation was significant for nominal groups ($F(1, 38) = 8.19$, $p = .007$), but not for interactive groups ($F < 1$). In line with previous research, this pattern of results was exactly the same as that found for productivity.

*Correlations.* The nonsignificant overall difference in quality between the generated and the selected ideas suggests that my participants' selection was not really better than chance, and that they would have done just as well taking a random sample from their production. If so, one would expect the quality of the selected ideas not to be dependent on the availability of good ideas (i.e., the productivity), but on the *average* quality of the generated ideas. This hypothesis was supported by the correlations between my dependent variables. Productivity was neither related to the originality ($r = -.13$, $p = .41$) nor to the feasibility of the selected ideas ($r = -.19$, $p = .23$), whereas the average originality of the generated ideas was related to the average originality of the selection ($r = .32$, $p = .036$), and the average feasibility of the generated ideas was related to the average feasibility of the selected ideas ($r = .41$, $p = .007$). Furthermore, and contrary to what would be expected on the basis of the brainstorming assumptions, the number of generated high-quality ideas was not related to the originality of the selected ideas ($r = .02$, $p = .93$) or the feasibility of the selected ideas ($r = .09$, $p = .56$). Thus, groups that generated more high-quality ideas did not select better ideas.

*Video Recordings*

Because participants in the one-task conditions were not explicitly instructed to separate idea generation and idea selection, it was possible that these participants would shift back and forth between the two tasks; such a strategy might affect both the generation and the selection process. Alternatively, they might spontaneously separate the tasks in the same way as the two-task participants were instructed to do: generate ideas first, and only start selecting ideas after the generation process had ended. Inspection of the video recordings showed that all of the one-task groups chose the latter option. In all groups, participants spent the first part of the task on idea generation exclusively. The onset of the selection process was usually marked by utterances such as "I don't think we can come up with more ideas; let's select the best ones" (in the interactive groups) and by the group members reaching for the colored selection form and starting to leaf through the ideas (in both interactive and nominal groups).

*Satisfaction*

Table 2.2 presents the means and standard deviation for the satisfaction scores; these were aggregated to the group level by averaging the three group members' scores. My participants were generally satisfied with their performance, both with regard to idea generation and idea selection: all overall means were significantly above the central value ($p < .001$). There was no difference between conditions in the satisfaction with the number of generated ideas ($p > .5$). Members of interactive groups were marginally more satisfied with the quality of their ideas ($M = 6.63$, $SD = 0.87$) than members of nominal groups ($M = 6.15$, $SD = 0.88$; $F(1, 38) = 3.13$, $p = .085$, partial $\eta^2 = .08$). Members of interactive groups were also more satisfied with the selection process ($M = 7.45$, $SD = 0.41$; $F(1, 38) = 7.39$, $p = .01$, partial $\eta^2 = .16$) and the quality of the selected ideas ($M = 7.38$, $SD = 0.77$; $F(1, 38) = 8.26$, $p = .007$, partial $\eta^2 = .18$) than members of nominal groups ($M = 7.0$, $SD = 0.69$, and $M = 6.62$, $SD = 0.90$, respectively). There were no other significant effects.

Table 2.2 - Means and Standard Deviations for Satisfaction in Study 2.1

| | 1 Task | | 2 Tasks | |
|---|---|---|---|---|
| Satisfaction measure | Nominal group | Interactive group | Nominal group | Interactive group |
| Number of generated ideas | 6.07 (0.84) | 6.20 (0.82) | 5.90 (1.78) | 5.94 (1.04) |
| Quality of generated ideas | 5.93 (0.97) | 6.60 (0.64) | 6.37 (0.76) | 6.67 (1.05) |
| Quality of selection process | 6.73 (0.62) | 7.47 (0.32) | 7.27 (0.68) | 7.44 (0.48) |
| Quality of selected ideas | 6.50 (0.80) | 7.10 (0.61) | 6.73 (0.94) | 7.61 (0.84) |

*Note*. Maximum value = 9. $N = 42$ groups. Standard deviations are in parentheses. All four items refer to participants' reported satisfaction, aggregated to the group level.

### *2.2.4 Discussion*

With regard to idea generation, my study replicated the production loss usually encountered in interactive brainstorming groups: Nominal groups generated more ideas, and more high-quality ideas, than interactive groups. This advantage of nominal groups was even more pronounced when the strict separation between idea generation and idea selection was removed. Interactive groups generated fewer ideas per minute in the one-task condition than in the two-task condition, which was in line with my prediction that removing the task separation would increase evaluation apprehension and production blocking in groups. However, because one-task interactive groups devoted more time to idea generation, they eventually were as productive as the two-task interactive groups. One-task nominal groups generated slightly more ideas per minute than two-task nominal groups. Because they were also more persistent, one-task nominal groups generated more ideas than did two-task nominal groups. Apparently, removing task separation increased participants' motivation to generate many ideas, perhaps because it was more salient that eventually they would need many good ideas to choose from. In interactive groups, evaluation apprehension and blocking counteracted higher productivity, but in nominal groups the increased motivation did lead to higher levels of performance. This finding qualifies the often-assumed benefits of task separation: Only in interactive groups did task separation lead to a more efficient brainstorming session (in terms of ideas per minute).

Although there were differences in the quality of the generated ideas (nominal groups generated ideas that were more original, interactive groups generated ideas that were more feasible; nominal groups also generated a larger number of high-quality ideas), these differences disappeared in the selection stage. It appears that interactive groups indeed managed to overcome their productivity loss by making an effective selection. However, the quality of the selected ideas was only slightly different from the average quality of the available ideas. The ideas selected by interactive groups were only marginally more original, and marginally less feasible, than their average production; the nominal groups did not even show this small difference. These findings suggest that, with regard to creativity, the selection process was hardly more effective than taking a random sample from the available ideas. This interpretation is supported by the pattern of correlations found, which shows that the quality of selected ideas was related to the average quality, and not the quantity, of the available ideas. Importantly, there was also no trace of a correlation between the number of generated good ideas and the quality of the selected ideas. Thus, productivity in this study does not seem to have done anything for the quality of the selected ideas, which goes directly against the quantity-breeds-quality argument behind the brainstorming procedure.

A possible explanation of the low selection effectiveness found in this study is that the originality and feasibility of the available ideas may have been strongly negatively correlated, making it difficult for participants to select good ideas. I therefore computed the average correlation between originality and feasibility for the generated ideas and for the selected ideas[4] (first performing a Fisher $r$-to-$Z$ transformation on all correlations, and then transforming the average value back to $r$). The average correlation between originality and feasibility was only marginally significant for the generated ideas ($r = -.28$, $p = .07$) and not significant for the selected ideas ($r = -.10$, $p = .60$). Thus, while there was a slight negative association of originality and feasibility in the available ideas, this correlation was not very strong.

My participants' low selection effectiveness notwithstanding, they were generally satisfied with the quality of their selection. This raises the question of what criteria my participants used when selecting ideas. The experimental task was explicitly framed as a

[4] In some cases, these correlations could not be computed, due to the absence of variance in the quality scores. Thus, for the generated ideas, $n = 41$; for the selected ideas, $n = 25$.

creativity task, the purpose of which was to come up with creative ideas, but I did not explicitly instruct participants to use originality and feasibility as their selection criteria; thus, I cannot exclude the possibility that they used other criteria. For example, because the brainstorming topic I used was highly relevant to my participants, it is plausible that they simply selected those ideas that they found most important, rather than the most original or feasible ideas. This view is supported by inspection of the video recordings: Many of the group discussions revolved around specific complaints regarding the Psychology curriculum, teachers and classes. Thus, it is possible that the poor selection performance found in this study was due to the lack of specific selection criteria in the experimental instructions, and that my participants would in fact have been able to make a more creative selection, had they been provided with more explicit selection criteria. Testing this hypothesis was the aim of Study 2.2.

## 2.3 Study 2.2: Idea Selection With Selection Criteria

### *2.3.1 Introduction*

In order to test the hypothesis that giving participants explicit selection criteria would improve the selection of creative ideas, I conducted a small follow-up study. In this experiment, I presented participants with a randomly chosen set of ideas from Study 2.1, and asked them to select the four best ideas from this set. I provided half of the participants with explicit selection criteria; the other participants were simply instructed to select "the best" ideas. If the poor selection performance in Study 2.1 was caused by a lack of explicit selection criteria, providing participants with such selection criteria should improve selection performance.

### *2.3.2 Method*

*Participants*

Thirty undergraduate Psychology students (22 females and 8 males, mean age = 21.2 years) participated in the experiment for course credit.

*Independent Variables and Materials*

This study involved one independent variable (instructions) with two levels (criteria vs. no criteria). Participants were assigned randomly to one of the experimental conditions.

I randomly selected 18 idea sets from those generated by individuals in the first experiment and distributed these sets randomly across the two conditions (12 of the sets were used in both conditions). Every participants received one idea set.

*Procedure*

In this study, all participants worked individually. Upon their arrival in the lab, participants read and signed an informed-consent form. They then received the first instruction. In the instruction, it was explained that they would receive a set of ideas; it was also explained that these ideas were generated by another undergraduate Psychology student in a brainstorming session, and that it would be their task to select the four best ideas from this set. Participants in the criteria condition were instructed to make this selection on the basis of the originality and feasibility of the ideas; participants in the no criteria condition were simply instructed to select “the best four ideas” without further explanation. Participants then received a list of ideas, and a selection sheet. These selection sheets were identical to those in Study 2.1, except for the selection instructions (criteria vs. no criteria), which were repeated at the top of the sheet.

After making their selection, participants filled in a questionnaire; they were then debriefed and thanked for their participation.

*Measures and Dependent Variables*

*Manipulation checks.* In the exit questionnaire, the participant was asked to rate the degree to which she or he used originality and feasibility as a criterion in making her or his selection. Exploratively, I also asked participants whether they had selected ideas that they themselves would like to see implemented, and whether they had selected ideas that were related to specific problems they themselves had encountered. Participants could respond to these four items on a 5-point scale (1 = *not at all*; 5 = *very much*).

*Idea quality.* For all ideas, the originality and feasibility ratings from Study 2.1 were used. Similarly to Study 2.1, I computed the mean originality and feasibility of the three favorite ideas selected by each participant. Because all participants in this study worked individually, and I was specifically interested in comparing participants’ performance in this study to participants’ performance in Study 2.1, I analyzed the data at the level of the individual.

*2.3.3 Results*

*Manipulation Checks*

I performed separate one-way ANOVAs on my questionnaire items. Participants in the criteria condition reported having relied on originality as a selection criterion more strongly ($M$ = 3.79, $SD$ = 1.12) than participants in the no criteria condition ($M$ = 1.67, $SD$ = 0.48; $F(1, 27) = 44.59$, $p < .001$, partial $\eta^2 = .62$). Participants in the criteria condition also reported having relied on feasibility as a criterion more strongly ($M$ = 4.33, $SD$ = 0.62) than participants in the no criteria condition ($M$ = 3.60, $SD$ = 1.29), although this difference was only marginally significant ($F(1, 28) = 3.90$, $p = .058$, partial $\eta^2 = .12$). Apparently, my manipulation was successful, although it appears that feasibility is a criterion people tend to use spontaneously in idea selection.

Interestingly, I found no difference between conditions on the other questionnaire items. Participants in both conditions reported having selected ideas that concerned specific problems they had encountered ($M$ = 3.48, $SD$ = 1.21, $F < 1$) and having selected ideas that they would like to see implemented ($M$ = 4.33, $SD$ = 0.71, $F < 1$).

*Idea Quality*

A 2 (Originality: originality of the idea set vs. originality of the selected ideas) x 2 (Instructions: criteria vs. no criteria) mixed model ANOVA yielded no significant effects ($F$s < 1). Thus, the average originality of the selected ideas was not significantly different from the average originality of the ideas in the idea set, and this was the case in both conditions.

A 2 (Feasibility: feasibility of the idea set vs. feasibility of the selected ideas) x 2 (Instructions: criteria vs. no criteria) mixed model ANOVA yielded no significant effects ($F$s < 1). Thus, as with originality, the average feasibility of the selected ideas was not significantly different from the average feasibility of the ideas in the idea set, and this was the case in both conditions.

### *2.3.4 Discussion*

This study was designed to test the effects of giving participants explicit selection criteria. I found no effect of criteria: Even though participants in the criteria condition reported using the criteria originality and feasibility to a stronger degree than participants in the no criteria condition, this made no difference for the quality of their selection. This makes it unlikely that the poor selection performance found in Study 2.1 was due to the absence of specific selection criteria. Nevertheless, because all participants in Study 2.2 worked individually, I cannot exclude the possibility that providing interactive groups with explicit selection criteria would enhance selection performance.

An unexpected finding is that providing participants with selection criteria did not decrease their reliance on other criteria. Specifically, all participants reported having selected ideas that concerned problems they had encountered, and ideas that they would like to see implemented, regardless of whether they were instructed to select original and feasible ideas.

## 2.4 General Discussion

Previous brainstorming research has focused on idea generation, but largely neglected idea selection. The primary aim of Study 2.1 was to test the implicit assumption that the higher productivity of nominal groups, which leads to a higher availability of high-quality ideas, also leads to their selecting better ideas. My results show that productivity is not enough: Although nominal groups generated more ideas, and more good ideas, than interactive groups, they did not select better ideas. In other words: the lower productivity of interactive groups did not put them at a disadvantage when it came to idea selection. However, in all conditions, the effectiveness of the selection was very low, as witnessed by the fact that the average quality of the selected ideas was not significantly different from the average quality of the generated ideas (or at best only marginally so). Thus, it cannot really be said that interactive groups made a *good* selection.

I hypothesized that the poor selection performance in Study 2.1 could have been due to a lack of specific selection criteria. Study 2.2 was designed to test this hypothesis. However, results showed that participants who were instructed to select ideas that were original and feasible did not perform better than participants who were simply instructed to select the best ideas. It would seem, therefore, that the results of Study 2.1. reflect something more fundamental in the way my participants selected ideas. (Interestingly, the fact that my participants appear to be unable to distinguish good from poor ideas is consistent with findings reported by Simonton (e.g., 2003) in his analysis of creativity in science and literature. Simonton reports that people are not very good at recognizing their best ideas, and that they do not improve over the course of their career.) The fact that all participants in Study 2.2 reported a tendency to select ideas that concerned specific problems they had encountered that they would like to see implemented, hints at an underlying preference for *solutions*, rather than *innovations*. This may be due to the nature of the brainstorming problem used in these studies: Most of my participants were undergraduate Psychology students, and they generated ideas about possible improvements in the Psychology curriculum. Their strong personal involvement may have decreased their tendency to select creative ideas (e.g., Harari & Graham, 1975). It is possible, then, that my results would have been different, had I used another brainstorming topic. The studies reported in Chapter 3, therefore, employ a different brainstorming topic.

Importantly, my results go directly against the assumption that a higher availability of high-quality ideas leads to the selection of better ideas. I found no evidence of such a relation: The higher productivity of nominal groups did not lead to their selecting better ideas, and there was no overall correlation between the number of generated good ideas and the quality of the selected ideas. Instead, measures of *average* originality and feasibility were significantly related to the quality of the selected ideas. Thus, while previous studies have focused on the *absolute* availability of good ideas, my results suggest that, given the chance-level selection performance that I found, the average quality of the available ideas is a better predictor of the quality of selected ideas.

It should be noted that, in Study 2.1, nominal group members generated *and* selected ideas individually. Thus, none of them had access to the total group's production. This may have contributed to the nominal groups' failure to profit from their high productivity. However, given the low effectiveness of the selection process, I suspect that pooling the generated ideas before the selection task would not have resulted in a higher quality of selected ideas. If people are not very good at recognizing and selecting their best ideas, it cannot be expected that selecting from the nominal group's total production will lead to a higher quality choice.

The results of these two studies raise some serious doubts with regard to brainstorming. If generating more ideas does not lead to the selection of better ideas, as my results indicate, in what sense does "quantity breed quality," as the brainstorming procedure presupposes? What exactly is the relationship between productivity and idea quality? This question will be addressed in Chapter 3, where I will test two alternative views of the quantity-quality relationship. Then, in Chapter 4, I will go deeper into the issue of selection effectiveness, and the question how selection effectiveness can be improved.

CHAPTER 3

# Knowledge Accessibility and Creative Performance [1]

## 3.1 Introduction

As explained in the previous chapters, the principle of *quantity breeds quality* states that the more ideas are generated, the more good ideas will be found among them. It is useful to generate large numbers of ideas, because this increases the availability of high-quality ideas, and this in turn increases the chance that good ideas will subsequently be selected for further development and implementation. Although the studies reported in the previous chapter cast some doubt on this latter expectation, there seems little doubt that enhancing productivity at least enhances the *potential* for creative idea selection.

Given the widespread use of brainstorming and related idea generation techniques, it is surprising that not much is known about the nature of the quantity-quality relationship. Because prior brainstorming studies have focused almost exclusively on the quantity of ideas generated during brainstorming, relatively little attention has been paid to the *quality* of ideas generated in a brainstorming session, let alone to the way in which quantity leads to quality. The brainstorming studies that have taken idea quality into account, have shown that the number of original and feasible ideas is strongly correlated with the total number of generated ideas. For example, Parnes and Meadow (1959) found a strong positive correlation between the total number of generated ideas and the number of ideas that were judged by a rater to be both highly original and highly useful. Similarly, Diehl and Stroebe (1987) found, both in a review of the literature and in an experiment, that the higher productivity of nominal groups as compared to interactive groups was consistently accompanied by the production of more highly original and feasible ideas. Indeed, the strength of this association is such that, as Diehl and Stroebe (1987) noted, little information seems to be gained from analyzing data on idea quality. However, in this chapter I will argue that the relationship between quantity and quality is more complex, and that data on idea quality can show patterns that differ significantly from general productivity. Specifically, I will argue that, contrary to what is commonly found, there is a systematic relationship between productivity and the average originality of generated

[1] This chapter is based on Rietzschel, Nijstad, and Stroebe (2005a).

ideas, but that this relationship may only be detected when productivity and quality are measured on a more specific level than is usually done. I will report three experimental studies to support this argument, but I will first go more deeply into some perspectives on the quantity-quality relationship.

### *3.1.1 Stochastic Creativity*

While the available data show that quantity indeed usually is accompanied by quality, in the sense that more high-quality ideas are available when more ideas are generated, it is not clear why this is the case. The most parsimonious interpretation of the quantity-quality relationship is to assume a chance process: For any generated idea, there is an equal (or near-equal) chance of its being an original and feasible idea. Therefore, according to the laws of chance, as more ideas are generated, the number of high-quality ideas generated is expected to increase linearly. This line of reasoning fits with the *Darwinian* view of creativity (Campbell, 1960), which has recently been developed further by Simonton (e.g., 1997, 1999a, 1999b). The production of new ideas is assumed to be a stochastic process, and the value or success of new ideas is supposed to be unpredictable. Idea generation can, in the chance view, be compared to flipping a coin: The chance of tails coming up (resp. generating a good idea) is supposed to be fixed (although not necessarily the same for each individual) and independent of earlier coin flips (resp. generated ideas). Thus, according to this view, increasing one's total productivity is the best, or even the *only* way to boost the output of high-quality creative ideas or solutions: If one wants to have more coins land tails up, one will have to throw more often.

The chance approach to the quantity-quality relation obviously predicts a linear relation between the number of generated ideas and the number of good ideas contained therein. This prediction is supported by the strong correlations found between these variables (e.g., $r = .69$ in Parnes and Meadow, 1959, and $r = .82$ in Diehl and Stroebe, 1987, Experiment 1; inspection of my own data from Study 2.1 similarly revealed a correlation of .69). The chance approach further predicts that the proportion of generated high-quality ideas and the *average* quality of generated ideas are unrelated to productivity (because generating more ideas also means generating more *low*-quality ideas), which is also confirmed by the available data (e.g., Dennis, Valacich, Connolly, & Wynne, 1996; Diehl & Stroebe, 1987).

### *3.1.2 The Benefits of Deep Exploration for Creativity*

Although there are empirical data to support the chance approach, several creativity researchers have proposed that the probability of generating a creative idea is not necessarily fixed and independent of productivity. Instead, these researchers argue that this probability can be affected by the approach that people take towards a problem, and the way in which they utilize their relevant knowledge. Originality is sometimes defined as the relative infrequency of an idea: The more people express an idea, the less original the idea is. The ideas that are mentioned by most people presumably are easiest to generate, and do not need much cognitive effort or special strategies. For example, when generating uses for a brick, many people will initially suggest using it to build something, such as a garden shed or a road. This is a very unoriginal application of a brick, but it is the first thing that comes to mind. Thus, the idea 'brick-as-building-block' has a much higher chance of being generated than other uses, such as the idea 'brick-as-percussion-instrument'.

*The Path-of-Least-Resistance Model*

Finke, Ward and Smith (1992) argue that people normally tend to rely on conventional lines of thought, unless they are actively encouraged to engage in deeper exploration of a problem. In line with this hypothesis, Ward (1994) and Ward, Patterson, Sifonis et al. (2002) found that participants had a strong tendency to rely on highly accessible properties and exemplars when generating new instances of specific categories (such as animals or tools), and this tendency limited the originality of participants' production. To explain these findings, Ward (1994) formulated the *path-of-least-resistance model*, which states that people tend to generate instances (or, in the present context, ideas) with the least cognitive effort possible. In many cases, this amounts to reproducing slightly modified (or even unmodified) instances that can be directly retrieved from memory.

Finke, Ward, and Smith (1992) propose that people might be induced to undertake deeper exploration of their knowledge about a particular problem if they are *restricted* in the elements they can use in generating an idea. Because they cannot rely on their normal, habitual ways of approaching the problem, people would then be forced to be more creative. For example, Finke et al. (1992) suggest that restricting the ways in which a particular problem is interpreted can force people to think beyond the normal options and to generate more "unusual possibilities and implications, which might not have been considered had the person been free to use any interpretive category" (p. 32). Thus, if people are forced to take another cognitive 'path' than the path of least resistance, the originality of their ideas or solutions should go up.

*3.1.3 Deep Exploration in Brainstorming*

In line with the path-of-least-resistance model, several brainstorming studies have found the originality of the generated ideas to be moderate or even low (e.g., Diehl & Stroebe, 1987; Putman & Paulus, 2003; Chapter 2, this dissertation), and there is evidence suggesting that brainstormers in organizations find it hard to overcome their inhibitions and be really original (Rickards, 1975). The notion that people who are required to generate new instances or ideas tend to simply reproduce highly accessible, conventional instances, fits well with my observations from Study 2.1, where I found that the group discussions in the idea selection stage tended to concern specific complaints that my participants had (the brainstorming topic was "possible improvements in education at the Psychology department" and my participants were undergraduate Psychology students). Presumably these complaints were highly accessible and could be retrieved very easily. Because all participants apparently had more or less the same complaints and generated ideas that were very similar, the generated ideas were not very original.

Further, Stein (1975) and Perkins (1981) both argue that people usually start out by generating conventional ideas, because these are founded upon highly accessible knowledge items. Only after these ideas have been verbalized and have thereby been 'removed' from the pool of potential ideas, or if the individual is motivated to keep searching for more unusual ideas (i.e., avoiding premature closure), will original and unconventional ideas be generated. In fact, this is exactly the reasoning underlying the quantity-breeds-quality argument put forward by Osborn (1963). Little empirical data on this issue are available, but Manske and Davis (1968) did find that the originality of ideas generated by their participants significantly increased over time. This is exactly what Osborn (1963) predicted, and cannot be explained by a chance interpretation.

Thus, there is some support in the brainstorming literature for both the path-of-least-resistance model and the deep exploration view of creativity. This raises the question of whether cognitive restrictions such as those proposed by Finke et al (1992) can enhance the generation of original ideas. At a first glance, this notion seems to be at odds with the atmosphere of freedom that the brainstorming rules aim to establish, but this is not really the case. The restrictions that Finke et al. refer to are not motivational or evaluative in nature, but rather concern the way in which a particular problem is approached during idea generation. Without restrictions, people may not be very effective in exploring their own knowledge about a brainstorming topic, and therefore be less creative. With restrictions, people may be compelled to look beyond their routine idea generation strategies and be truly creative.

*Subcategories and Deep Exploration*

Dennis, Valacich, Connolly, and Wynne (1996) report a study in which restrictions indeed did lead to better brainstorming performance. They hypothesized that brainstormers often limit themselves to exploring a few subcategories of the overall problem, overlooking possible solutions from other subcategories. Decomposing a brainstorming problem into separate subcategories, and explicitly instructing participants to generate ideas sequentially within each of these subcategories, should help them to overcome this self-limiting tendency and to generate more ideas, and more good ideas, and Dennis et al.'s results confirmed this hypothesis. Similarly, Coskun, Paulus, Brown, and Sherwood (2000) found that problem decomposition increased productivity in interactive and nominal groups.

Nijstad, Stroebe, and Lodewijkx (2002) performed a study in which participants were provided with stimulation ideas in a brainstorming task. These stimulation ideas either came from a wide range of semantic categories (heterogeneous stimulation), or from a limited number of semantic categories (homogeneous stimulation). Nijstad et al. found that participants who were provided with homogeneous ideas during brainstorming, generated as many ideas as participants who were provided with heterogeneous stimulation ideas, but that these ideas came from fewer subcategories of the brainstorming topic. In other words, homogeneous stimulation ideas decreased the number of categories that participants used, but increased the average number of ideas generated within subcategories. Thus, homogeneous stimulation ideas seemed to induce deeper exploration of subcategories than heterogeneous stimulation.

The deep exploration hypothesis explained above suggests that deeper exploration of a subcategory should not only be associated with higher productivity within that subcategory, but also with a higher originality of the ideas generated within that subcategory. Because there are fewer conventional ideas associated with each subcategory, deeper exploration of a subcategory should cause brainstormers to generate more unconventional, and hence original, ideas within that subcategory. However, Coskun et al. (2000) and Nijstad et al.(2002) only report productivity data, and Dennis et al. (1996) report only *global* measures of idea quality; that is, quality was summed and averaged across *all* generated ideas. Hence, this implication of the deep exploration hypothesis remains untested. I argue that it is vital to also examine productivity *and* idea quality within subcategories. Global measures of idea quantity and quality may well obscure patterns of creative performance that can be revealed by a more content-based analysis.

The studies cited above show that both problem decomposition and stimulation ideas can lead to deeper exploration of subcategories, but provide no data with regard to originality. The studies reported in this chapter aim to fill this void and to provide a test of the hypothesis that deep exploration can increase the originality of generated ideas.

*Priming and Accessibility*

Although problem decomposition and stimulation ideas have both been found to work, these manipulations have the disadvantage that they actually change the properties of the brainstorming task. Participants who are brainstorming about a decomposed problem are in effect performing a different task than participants for whom the problem is not decomposed, and participants who are provided with stimulation ideas receive other information during idea generation than participants who are not provided with stimulation ideas. Thus, neither of these manipulations is completely free from possible confounds. In order to assess the 'pure' effects of deep exploration, participants in a brainstorming task should be induced to engage in deep exploration without changing the nature of the brainstorming task itself. I hypothesized that this could be achieved through a *priming* manipulation.

Essentially, the effect of a priming manipulation is to increase the accessibility of particular constructs or knowledge. While a distinction is usually made between *chronic* and *temporary* accessibility, research in social cognition has shown that these are functionally similar in a number of respects (Bargh, Bond, Lombardi, & Tota, 1986; Higgins, 1989). Most importantly, both chronic and temporary sources of accessibility affect the "readiness with which available constructs are actually used" (Higgins, 1989, p. 82). Research on priming effects has shown that it is possible to increase temporary knowledge accessibility with relatively small-scale manipulations, such as subliminally flashing a concept on a computer screen (see Higgins, 1996 for a review). If the role of knowledge accessibility is as important as the path-of-least-resistance model suggests, priming could be an effective way to influence participants' idea generation without changing the nature of the brainstorming task itself.

In this chapter, I used a priming manipulation, as well as stimulation ideas, to increase the accessibility of participants' knowledge about a subcategory of a subsequent brainstorming problem. I expected this higher accessibility to lead to their generating relatively more ideas within that subcategory. In line with the deep exploration hypothesis, I further expected this higher accessibility to lead to their generating (on average) more original ideas within that subcategory. However, in line with my statement that global measures of idea quantity and quality are not sensitive to effects within subcategories, I expected no effects of knowledge accessibility on global productivity or global idea quality. Furthermore, because brainstorming in interactive groups remains highly popular, and because interaction creates the opportunity for mutual stimulation (Nijstad et al., 2002), I also tested the effect of knowledge activation within an interactive (dyadic) context.

## 3.2 Study 3.1: Effects of Priming on Idea Generation

### *3.2.1 Introduction*

In order to study the effects of knowledge activation on idea generation, I conducted an experiment in which participants answered open-ended questions prior to generating ideas in an individual brainstorming session. These questions, which were intended as a priming-type manipulation, concerned a sub-topic of the broader brainstorming topic. Specifically, the brainstorming topic concerned ways to improve or maintain one's health, and the priming questions concerned nutrition, hygiene, or sports. I employed two control conditions: one in which participants were not primed at all, and one in which participants were primed with an irrelevant topic (politics), that did not constitute a subcategory of the brainstorming problem. If I had only used a control condition without priming, effects of my manipulation might plausibly be attributed to a general activation of my participants' knowledge or cognitive effort, rather than to the specific *content* of my priming manipulation.

I hypothesized that my priming manipulation would induce participants to engage in deeper idea exploration of the primed sub-topic. The enhanced accessibility of relevant knowledge within a particular (primed) domain should increase the chance that generated ideas fall within the primed topic, and should also increase the chance that those generated ideas are relatively unusual (and therefore relatively original). Thus, I expected two specific results of this deeper processing. Firstly, I expected participants who had been primed with a specific sub-topic to generate a higher percentage of ideas within that sub-topic than participants who had not been primed. Secondly, I expected participants who had been primed with a particular sub-topic to generate ideas of higher originality within that sub-topic than participants who had not been primed.

### *3.2.2 Method*

*Participants*

Participants were 93 undergraduate students of Psychology at the University of Amsterdam.

*Design and Independent Variables*

The study employed one independent variable (Priming: nutrition, hygiene, sports, irrelevant priming about politics, or no priming). Participants were assigned randomly to one of the conditions.

*Priming questions.* All participants, except the participants in the no priming condition, were presented with a series of four open-ended questions. They could use a maximum of 200 words to answer each single question. These questions were as follows: (1) "How much time and attention do you usually devote to healthy nutrition [hygiene, sports, politics]? Please provide some examples of your behavior." (2) "Do you usually devote more or less time to healthy nutrition [hygiene, sports, politics] than other people? Please provide some examples of why you think this is the case." (3) "Do you think you ought to devote more time and attention to healthy nutrition [hygiene, sports, politics]? Please provide some examples of why you think this is the case." (4) "How important do you think it is for people to devote time and energy to healthy nutrition [hygiene sports, politics]? Please explain why you think so."

*Procedure*

Participants were seated individually in front of an Apple iMac computer. When the experimenter had started the program, participants first read a general introduction to the experiment. In this introduction, they were told that they were about to participate in an experiment regarding the effect of elaborate cognitive processing on attitudes. The reason for this cover story was that I did not want my participants to suspect a link between the priming manipulation and the brainstorming session. Thus, participants (except the participants in the no priming condition, who went immediately to the idea generation task) were presented with three statements supposedly measuring their attitudes towards (depending on condition) nutrition, hygiene, sports, or politics.

After responding to these statements, participants were presented with the four priming questions, which were followed by a second "attitude measurement". This was followed by the idea generation task. Participants were told that they were to generate ideas in a brainstorming session, and that the goal of a brainstorming session is to come up with a large number of creative ideas. At this point, the four brainstorming rules were explained (quantity is wanted; freewheeling is welcomed; combine and improve ideas;

criticism is ruled out). Participants were informed that ten minutes would be available for idea generation. They were further instructed to enter all ideas that they could think of, but not to enter anything when they could not think of any ideas. Participants were then informed that the topic of the brainstorming session would be "what can people do to improve or maintain their health?" and could start generating ideas. After ten minutes had passed, participants were debriefed, thanked, and paid.

*Measures and Dependent Variables*

*Productivity.* Productivity was defined as the total number of unique ideas generated by each participant.

*Ideas Within Primed Topics.* All ideas were coded for content according to a coding system developed by Diehl (1991). This system consists of a matrix of eleven specific goals (e.g., "avoiding nutritional deficits") that are crossed with nine specific means (e.g., "food intake"), yielding 99 different possible categories (see Appendix 2). Each idea was assigned a numerical code that reflected a specific category. A second rater coded a randomly selected subset of 150 ideas. The raters assigned the same code in 84 % of the cases. In light of the large number of possible categories (and hence the small probability of chance agreement), I deemed this to be sufficiently reliable and used the first rater's codes. These codes were then used to classify ideas according to content, which allowed me to measure whether an idea fell within the domain of nutrition, hygiene, or sports. The coding system used explicitly distinguishes between different topics. Thus, all ideas for which the means was coded as "food intake" and all ideas for which the goal was coded as "avoiding nutritional deficits" were initially coded as nutrition ideas. Similarly, all ideas for which the means was coded as "physical exercise" and all ideas for which the goal was coded as "improving physical fitness" were initially coded as sports ideas. Further, all ideas for which the means was coded as "physical cleanliness" were initially coded as hygiene ideas. Next, all ideas (regardless of category code) were inspected for explicit references to nutrition, sports or hygiene and assigned to the appropriate categories.[2]

During the coding process, ideas were placed in randomized order and the coders were blind to conditions.

*Idea Quality.* All ideas were scored on a 5-point scale by a trained rater for originality and feasibility. A second rater scored a subset of 250 ideas. With raters

[2] Note that according to this procedure, ideas could theoretically be classified in more than one subcategory (e.g., nutrition *and* sports). Inspection of the assigned codes showed that only one idea was assigned to more than one subcategory; this was an idea suggesting the tearing down of all establishments of a well-known fast-food chain, because of the low quality of its food and the unhygienic circumstances. This idea was included in both categories.

considered to be in agreement whenever their ratings differed by no more than one point (Diehl & Stroebe, 1987), agreement existed in 98.8% of the cases for both originality and feasibility. I also computed the intraclass correlations (see Study 2.1); these were .75 for originality and .78 for feasibility, which was satisfactory. For each participant, I computed the mean originality and feasibility of the generated ideas, as well as the number of generated high-quality ideas (defined as ideas with originality and feasibility scores above the mean) as measures of *global* idea quality. I further computed the mean originality of the ideas generated *within* each of the three subcategories (nutrition, hygiene, and sports).

### *3.2.3 Results*

*Productivity*

*Global Productivity.* All means and standard deviations for productivity and the ideas within primed topics are presented in Table 3.1. A one-way ANOVA on productivity yielded no significant effect of my priming manipulation ($F$ < 1): As expected, my manipulation did not affect the number of ideas generated by my participants ($M$ = 16.91; $SD$ = 8.91).

*Productivity Within Primed Topics.* I had specific and directional hypotheses with regard to the percentages of ideas generated within primed topics, and the originality of those ideas. Hence, I performed one-tailed planned comparisons (contrast analyses) to test these hypotheses.

My hypotheses with regard to the percentages of ideas within primed topics were twofold. Firstly, because I did not expect irrelevant priming (politics) to increase the accessibility of knowledge within the subcategories nutrition, hygiene or sports, I expected participants with irrelevant priming to generate as many ideas within the three topics nutrition, hygiene and sports as participants without priming. Secondly, I expected participants who were primed with a particular topic to generate a higher percentage of ideas within that topic than participants in the two control conditions.

Planned comparisons for the three primed topics showed that participants with irrelevant priming generated the same percentages of nutrition ideas ($M$ = 21.51, $SD$ = 11.85), hygiene ideas ($M$ = 1.99, $SD$ = 3.24) and sports ideas ($M$ = 17.03, $SD$ = 11.94) as participants who were not primed ($M$ = 17.77, $SD$ = 9.03; $M$ = 2.32, $SD$ = 5.31; $M$ = 14.32, $SD$ = 7.97, for nutrition, hygiene and sports, respectively; all $t$s < 1). Thus, as expected, the two control conditions were identical with regard to the percentages of ideas within the three primed topics. Other effects of my priming manipulation can therefore not be attributed to general consequences of priming per se.

Planned comparisons further showed that participants with nutrition priming generated a higher percentage of nutrition ideas (*M* = 32.08, *SD* = 15.83) than participants in the two control conditions ($t(88) = 3.92$, $p < .001$, $r = .81$[3]), that participants with hygiene priming generated a higher percentage of hygiene ideas (*M* = 10.98, *SD* = 15.07) than participants in the two control conditions ($t(88) = 4.09$, $p < .001$, $r = .85$), and that participants with sports priming generated a higher percentage of sports ideas (*M* = 21.67, *SD* = 14.91) than participants in the two control conditions ($t(88) = 1.87$, $p = .033$, $r = .56$).

Thus, across the three topics, participants who were primed with a particular topic generated relatively more ideas within that topic than participants who were primed with an irrelevant topic or who were not primed at all.

Table 3.1 - Means and Standard Deviations for Study 3.1

| | Priming | | | | |
|---|---|---|---|---|---|
| Measure | Nutrition | Hygiene | Sports | Politics | No priming |
| Productivity | 16.29 (8.50) | 16.06 (6.51) | 18.88 (13.67) | 16.85 (7.73) | 16.71 (7.57) |
| Perc. of nutrition ideas | 32.08 (15.83) | 16.64 (9.18) | 18.43 (9.37) | 21.51 (11.85) | 17.77 (9.03) |
| Perc. of hygiene ideas | 1.91 (4.40) | 10.98 (15.07) | 0.35 (1.43) | 1.99 (3.24) | 2.32 (5.31) |
| Perc. of sports ideas | 19.92 (9.41) | 10.94 (9.12) | 21.67 (14.91) | 17.03 (11.94) | 14.32 (7.97) |
| Originality of ideas [a] | 1.65 (0.32) | 1.70 (0.27) | 1.78 (0.27) | 1.79 (0.34) | 1.68 (0.21) |
| Orig. of nutrition ideas [a b] | 1.44 (0.54) | 1.11 (0.22) | 1.26 (0.34) | 1.28 (0.47) | 1.19 (0.52) |
| Orig. of hygiene ideas [a c] | 2.30 (0.45) | 2.07 (0.26) | 1.67 (0.00) | 2.08 (0.20) | 2.00 (0.00) |
| Orig. of sports ideas [a b] | 1.27 (0.34) | 1.13 (0.28) | 1.50 (0.52) | 1.35 (0.41) | 1.21 (0.28) |

*Note.* *N* = 93 participants, except where otherwise indicated. Standard deviations are in parentheses.
[a] Maximum value = 5.
[b] *N* = 91.
[c] *N* = 27.

*Idea Quality*

*Global Idea Quality.* I performed a repeated measures ANOVA with the average originality and feasibility of the generated ideas as two levels of a within-subjects factor idea quality, and a between–subjects factor priming. This analysis yielded a main effect of idea quality: The average originality of the generated ideas (*M* = 1.72, *SD* 0.29) was lower than the average feasibility of the generated ideas (*M* = 4.36, *SD* = 0.34; $F(1, 88) = 1932.61$, $p < .0001$, partial $\eta^2 = .96$). There were no other significant effects, indicating

[3] The *r* I report for my contrast analyses is the $r_{\text{alerting}}$. This is an effect size measure that reflects the correlation between the observed group means and the contrast weights reflecting the predicted pattern of group means (Furr, 2004).

that there were no differences between conditions in global idea quality. In general, global originality and feasibility were strongly negatively correlated ($r = -.71, p < .001$).

An ANOVA with the number of generated high-quality ideas yielded no effect of priming ($F < 1$), which is similar to my finding with regard to global productivity.

*Originality of Ideas Within Primed Topics.* I hypothesized a relation between the percentage of ideas generated within a particular topic and the average originality of those ideas: As participants generated relatively more ideas within a particular topic, I expected the average originality of those ideas to be higher. As an initial test of this hypothesis, I computed correlations between the percentages of ideas generated within a topic (i.e., nutrition, hygiene or sports) and the average originality of those ideas. I found a positive correlation between participants' percentage of nutrition ideas and the originality of the nutrition ideas ($r = .36, p < .001$), and a positive correlation between participants' percentage of sports ideas and the originality of the sports ideas ($r = .61, p < .001$). However, no such correlation was found for the hygiene ideas ($r = .04, p > .8$). This was probably due to the large number of missing cases: Since 66 of the 93 participants generated no hygiene ideas at all, and no originality measures could therefore be computed for these participants, variance and statistical power were limited.

Thus, in general, as participants generated relatively more ideas within a particular topic, the average originality of these ideas increased. Importantly, global productivity was *not* correlated with global originality ($r = .08, p > .4$). Hence, it was not so much global productivity that enhanced originality, as the *relative* productivity *within* a particular topic.

Next, I performed planned comparisons similar to those reported above. If generating relatively many ideas within a particular topic indeed is associated with a higher originality of those ideas, contrast analyses should yield comparable results.

As expected, planned comparisons showed that the originality of the generated nutrition, hygiene and sports ideas was equally high for participants with irrelevant priming ($M = 1.28$, $SD = 0.47$; $M = 2.08$, $SD = 0.20$; $M = 1.35$, $SD =0.41$, for nutrition, hygiene and sports, respectively) as for participants who were not primed ($M = 1.19$, $SD = 0.52$; $M = 2.00$, $SD = 0.00$; $M = 1.21$, $SD = 0.28$; $t < 1$ for nutrition and hygiene, $t(88) = 1.13$, $p = .26$ for sports).

Planned comparisons further showed that the nutrition ideas generated by participants with nutrition priming were more original ($M = 1.44$, $SD = 0.54$) than those generated by participants in the two control conditions ($t(86) = 1.74$, $p = .043$, $r = .71$), and that the sports ideas generated by participants with sports priming were more original ($M = 1.50$, $SD = 0.52$) than those generated by participants in the two control conditions ($t(86) = 1.98$, $p = .026$, $r = .65$). However, the hygiene ideas generated by participants with hygiene priming were not more original ($M$ = 2.07, $SD$ = 0.26) than those generated by participants in the two control conditions ($t(22) = 0.243$, $p = .41$).

These results largely support my hypothesis that the generation of a relatively high percentage of ideas within a particular topic is associated with higher originality for those ideas, and that inducing participants to generate a higher percentage of ideas within a topic can increase the originality of these ideas.[4]

### *3.2.4 Discussion*

My results show that my priming manipulation was successful. Instructing participants to think elaborately about a particular subcategory of the later brainstorming topic did lead to their generating relatively more ideas within that subcategory, presumably through a relatively higher activation of the relevant knowledge structures. My results also show that generating more ideas within a particular subcategory was associated with a higher originality for those ideas, which supports the deep exploration hypothesis. However, this did not affect the overall quality of the ideas, nor did it affect overall productivity. Thus, while participants with priming apparently engaged in deeper exploration of a particular subcategory of the overall brainstorming problem, this did not affect their overall performance. Furthermore, global productivity and global originality were not correlated, while productivity and originality *within* the subcategories nutrition and sports were significantly correlated. These findings show that global measures of idea quantity and quality can indeed obscure important patterns in participants' creative performance.

## 3.3 Study 3.2: Effects of Priming and Stimulation Ideas

### *3.3.1 Introduction*

Considering the interactive nature of most real-world brainstorming sessions, an important question to study is what the effects of my priming manipulation would be in a face-to-face situation. Research on cognitive stimulation in brainstorming settings has shown that being exposed to other people's ideas can strongly influence one's own idea generation process (e.g., Dugosh, Paulus, Roland, & Yang, 2000; Nijstad et al., 2002). Therefore, it would be especially interesting and, given the widespread use of group

[4] I also tested these contrasts with the number of original ideas within the primed categories; in accordance with previous studies, these showed exactly the same pattern of means as productivity within subcategories. I defined an 'original idea' as an idea with an originality score above the average originality found in this study.

brainstorming, useful, to study the effects of knowledge activation in an interactive brainstorming situation.

*Stimulation Ideas*

As Nijstad et al. (2002) argue, testing hypotheses about the possible effects of being exposed to other people's ideas requires experimental control over the content of those ideas. Furthermore, a group context naturally affects creative performance in other ways than merely through cognitive stimulation or interference (e.g., Camacho & Paulus, 1995; Paulus & Dzindolet, 1993). Thus, before moving to an interactive brainstorming situation, I employed a cognitive stimulation paradigm (e.g., Dugosh et al., 2000; Nijstad et al., 2002) in combination with my priming manipulation. As in the previous study, participants in Study 3.2 were primed with questions about nutrition or hygiene, or were not primed at all. Then, during the brainstorming task, participants received stimulation ideas about nutrition or hygiene, or received no stimulation ideas at all. Stimulation ideas were presented on the computer screen, above the text field where participants could enter a new idea.

I expected both priming and stimulation to increase the relative accessibility of knowledge within the primed and stimulated categories, and therefore to increase the relative productivity and average originality within these categories. However, the *combination* of priming and stimulation was expected to be especially important. Nijstad et al. (2002) found that exposure to diverse ideas increased the *breadth* (i.e., the number of semantic categories used) of idea generation, while exposure to homogeneous ideas increased the *depth* (i.e., the average number of ideas within a semantic category) of idea generation. Because I argue that deeper idea generation is associated with a higher originality, I expected that participants for whom the priming questions and the stimulation ideas were homogeneous (i.e., concerned the same subcategory of the brainstorming topic) would engage in deeper exploration of that subcategory than participants for whom priming and stimulation were heterogeneous (i.e., for whom the priming questions and the stimulation ideas concerned different subcategories). Finally, in line with the findings of Dennis et al. (1996) that people apply their knowledge in idea generation rather superficially, I expected participants without priming and stimulation to explore the different subcategories least deeply. Thus, I expected participants with homogeneous priming and stimulation to show the highest productivity and average originality within subcategories, followed by participants with heterogeneous priming and stimulation, followed in turn by participants without priming and stimulation.

### *3.3.2 Method*

*Participants*

Participants were 94 undergraduate students at the University of Amsterdam (57 females, 37 males, mean age = 21,89 years), who participated for course credit or a reward of 7 Euros (about 9 US Dollars).

*Design and Independent Variables*

The study employed a 3 (Priming: nutrition, hygiene, or no priming) x 3 (Stimulation: nutrition, hygiene, or no stimulation) factorial design. Participants were randomly assigned to one of the conditions.

*Priming.* I used the same priming manipulation as in Study 3.1, but only used the topics nutrition and hygiene.

*Stimulation.* During brainstorming, participants were presented with stimulation ideas. These ideas had been taken from an earlier brainstorming experiment. The stimulation ideas concerned either nutrition or hygiene. The experimental program was set up in such a way that participants saw a new stimulation idea every time they entered an idea. A list of 30 stimulation ideas was available for both topics. Hence, if participants generated more than 30 ideas (which was the case for only seven participants), no new stimulation ideas were visible while they generated their last few ideas.

*Procedure*

The procedure of the study was very similar to that of Study 3.1, with the addition of the stimulation ideas. Participants were seated individually in front of an Apple iMac computer. As in Study 3.1, participants read a general introduction in which they were told that the experiment concerned the effect of elaborate cognitive processing on attitudes. Participants (except the participants in the no priming condition, who went immediately to the idea generation task) were then presented with an "attitude measurement," followed by the priming questions and a second "attitude measurement." This was followed by the actual idea generation task. As in the previous study, participants were told that they were to generate ideas in a brainstorming session, and that the goal of a brainstorming session is to come up with a large number of creative ideas. At this point, the four brainstorming rules were explained. Participants were informed that ten minutes would be available for idea generation. They were further instructed to enter all ideas that they could think of, but not to enter anything when they could not think of any ideas. Participants were then informed that the topic of the brainstorming session would be "what can people do to improve or maintain their health?" Furthermore, participants (except those in the no stimulation condition) were informed that during the brainstorming task, stimulation ideas would be visible on the computer screen. These ideas were presented above the text field where participants entered their ideas; after a participant pressed the Enter key to submit an idea, the screen was refreshed and a new stimulation idea was presented. Participants were instructed to pay attention to these stimulation ideas, because they could be helpful in generating ideas. Participants could then start generating ideas.

After the allotted ten minutes had passed, participants were thanked, debriefed and paid.

*Measures and Dependent Variables*

*Global Productivity.* Global productivity was defined as the number of unique ideas generated by each participant.

*Ideas Within Primed Topics.* I used the same procedure as in Study 3.1 to measure the percentages of ideas within the primed topics nutrition and hygiene. Thus, all ideas were assigned a 'goal x means' code; a second rater coded 250 ideas. The raters assigned the same code in 92.8 % of the cases. I deemed this to be sufficiently reliable and used the first rater's codes to further classify the ideas according to content.

*Idea Quality.* All ideas were scored on a 5-point scale for originality and feasibility by a trained rater. A second rater scored a subset of 250 ideas. With raters considered to be in agreement whenever their ratings differed by no more than one point, agreement existed in 100% of the cases for originality and in 98.4% of the cases for feasibility. Intraclass correlation coefficients were .78 for originality and .77 for feasibility. For each participant, I computed the average originality and feasibility of the generated ideas as measures of global quality. I further computed the average originality of the generated nutrition and hygiene ideas.

### *3.3.3 Results*

*Productivity*

*Global Productivity.* Means and standard deviations for the global variables reported in Study 3.2 are presented in Table 3.2. I performed a 3 (Priming) x 3 (Stimulation) ANOVA with productivity as the dependent variable and found no significant effects ($p$s $> .1$). Thus, neither priming nor stimulation affected the global productivity of my participants.

*Productivity Within Primed Topics.* My hypotheses with regard to productivity within topics concerned only the experimental conditions in which participants were primed and also received stimulation ideas, and the condition in which participants were not primed and received no stimulation ideas. Hence, I re-labeled the nutrition priming/nutrition stimulation condition as the homogeneous nutrition condition, and the hygiene priming/hygiene stimulation condition as the homogeneous hygiene condition. I further collapsed the two conditions nutrition priming/hygiene stimulation and hygiene priming/nutrition stimulation into one condition and labeled this the heterogeneous condition. The no priming/no stimulation condition was labeled the control condition. Table 3.3 presents the means and standard deviations of the percentages of ideas within subcategories for these re-labeled conditions.

Table 3.2 - Means and Standard Deviation for Study 3.2

| | *Nutrition priming* | | |
|---|---|---|---|
| | Stimulation ideas | | |
| Measure | Nutrition | Hygiene | None |
| Productivity | 19.22 (5.59) | 18.82 (7.19) | 20.17 (8.28) |
| Originality of ideas [a] | 2.08 (0.33) | 1.93 (0.25) | 1.71 (0.28) |
| | *Hygiene priming* | | |
| | Stimulation ideas | | |
| Measure | Nutrition | Hygiene | None |
| Productivity | 16.64 (6.57) | 22.63 (8.26) | 18.89 (12.90) |
| Originality of ideas [a] | 1.87 (0.44) | 1.89 (0.23) | 1.71 (0.13) |
| | *No priming* | | |
| | Stimulation ideas | | |
| Measure | Nutrition | Hygiene | None |
| Productivity | 19.36 (5.73) | 23.67 (4.33) | 14.73 (6.37) |
| Originality of ideas [a] | 1.64 (0.27) | 1.90 (0.19) | 1.71 (0.28) |

*Note.* $N$ = 94 participants. Standard deviations are in parentheses.
[a] Maximum value = 5.

*Nutrition Ideas.* I expected participants in the homogeneous nutrition condition to generate a higher percentage of nutrition ideas than participants in the heterogeneous condition and the control condition, and I expected participants in the heterogeneous condition to generate a higher percentage of nutrition ideas than participants in the control condition. Planned comparisons confirmed this hypothesis: Participants in the homogeneous nutrition condition generated a higher percentage of nutrition ideas ($M$ =

68.51, *SD* = 13.41) than participants in the heterogeneous condition (*M* = 31.69, *SD* = 20.57; $t(49) = 5.87$, $p < .001$, $r = .71$), and a higher percentage of nutrition ideas than participants in the control condition (*M* = 15.70, *SD* = 9.22; $t(49) = 7.28$, $p < .001$, $r = .88$). Participants in the heterogeneous condition also generated a higher percentage of nutrition ideas than participants in the control condition ($t(49) = 2.74$, $p = .005$, $r = .33$).

*Hygiene Ideas*. For the hygiene ideas, I expected a similar pattern as for the nutrition ideas: Participants in the homogeneous hygiene condition were expected to generate the highest percentage of hygiene ideas, followed by participants in the heterogeneous condition, followed in turn by participants in the control condition. Planned comparisons showed that, as expected, participants in the homogeneous hygiene condition generated a higher percentage of hygiene ideas (*M* = 29.65, *SD* = 11.88) than participants in the heterogeneous condition (*M* = 16.91, *SD* = 20.22; $t(49) = 2.09$, $p = .021$, $r = .44$) and the control condition (*M* = 2.85, *SD* = 4.33; $t(49) = 3.85$, $p < .001$, $r = .81$), and participants in the heterogeneous condition generated a higher percentage of hygiene ideas than participants in the control condition ($t(49) = 2.59$, $p = .007$, $r = .54$).

*Idea quality*

*Global Quality*. A mixed model ANOVA with the originality and feasibility of the generated ideas as two levels of a within-subjects factor quality, and priming and stimulation as between-subjects factors, yielded a main effect of quality ($F(1, 85) = 1849.06$, $p < .0001$, partial $\eta^2 = .96$), indicating that the generated ideas were, on average, less original (*M* = 1.82, *SD* = 0.32) than feasible (*M* = 4.49, *SD* = 0.33). There were no further main or interaction effects (*p*s > .1). As in the previous study, global originality and global feasibility were strongly negatively correlated ($r = -.73$, $p < .001$).

Table 3.3 - Means and Standard Deviations for Re-labeled Conditions in Study 3.2

| | Condition (re-labeled) | | | |
|---|---|---|---|---|
| Measure | Homogeneous nutrition | Homogeneous hygiene | Heterogeneous | Control |
| Percentage of nutrition ideas | 68.51 (13.41) | 12.48 (7.00) | 31.69 (20.57) | 15.70 (9.22) |
| Originality of nutrition ideas[a] | 2.10 (0.42) | 1.27 (0.26) | 1.64 (0.59) | 1.28 (0.41) |
| Percentage of hygiene idea | 0.62 (1.85) | 29.65 (11.88) | 16.91 (20.22) | 2.85 (4.33) |
| Originality of hygiene ideas[a] | 2.00 (0.00)[b] | 2.23 (0.37) | 2.06 (0.16) | 2.00 (0.00) |

*Note*. The heterogeneous condition contains the nutrition priming/hygiene stimulation and the hygiene priming/nutrition stimulation conditions.
[a] Maximum value = 5.
[b] *N* = 1.

*Originality of Nutrition Ideas*. Means and standard deviations for the originality of the ideas within the primed topics are presented in Table 3.3. I expected the originality of the ideas within the primed topics to show the same pattern of results as the percentages of ideas within the primed topics. Thus, I used the same re-labeled conditions as in the analyses reported above (homogeneous, heterogeneous and control) and tested the same contrasts (homogeneous against heterogeneous and against control; heterogeneous against control).

My planned comparisons showed that, as expected, the nutrition ideas generated by participants in the homogeneous nutrition condition ($M = 2.10$, $SD = 0.42$) were more original than those generated by participants in the heterogeneous condition ($M = 1.64$, $SD = 0.59$; $t(49) = 2.36$, $p = .011$, $r = .57$) and the control condition ($M = 1.28$, $SD = 0.41$; $t(49)= 3.66$, $p < .001$, $r = .89$), and the nutrition ideas generated by participants in the heterogeneous condition were more original than those generated by participants in the control condition ($t(49) = 1.98$, $p = .027$, $r = .48$).

*Originality of Hygiene Ideas*. Planned comparisons showed that the hygiene ideas generated by participants in the homogeneous hygiene condition ($M = 2.23$, $SD = 0.37$) were only marginally significantly more original than those generated by participants in the heterogeneous condition ($M = 2.06$, $SD = 0.16$; $t(22) = 1.58$, $p = .064$, $r = .84$) and the control condition ($M = 2.00$, $SD = 0.00$; $t(22) = 1.56$, $p = .066$, $r = .83$), and the hygiene ideas generated by the participants in the heterogeneous condition were not more original than those generated by participants in the control condition ($t < 1$).

These results largely support my hypotheses. Priming and stimulation caused higher productivity within the subcategories nutrition and hygiene, especially when priming and stimulation were homogeneous. For the nutrition ideas, this pattern of results was also followed by the originality of the ideas. For the hygiene ideas, this was only partly the case. However, in accordance with my hypothesis, participants in the homogeneous hygiene condition, who generated the highest percentage of hygiene ideas, also generated the most original hygiene ideas. Furthermore, it should be noted that, as in Study 3.1, a large number of participants (50 out of 93 participants) generated no hygiene ideas at all, resulting in limited variance and statistical power.

As in the previous study, the percentage of generated nutrition ideas was positively correlated with the average originality of the nutrition ideas ($r = .38$, $p < .001$). However, the percentage of hygiene ideas was not significantly related to the originality of the hygiene ideas ($r = .15$, $p = .33$). Furthermore, global productivity and global originality were not significantly correlated ($r = -.01$, $p = .95$), indicating that productivity in itself did not increase the average originality of the generated ideas.

### *3.3.4 Discussion*

In accordance with my hypotheses, priming questions and stimulation ideas affected the content and quality of the generated ideas, and this effect was strongest when priming and stimulation were homogeneous with regard to topic. As participants generated relatively more ideas within a particular topic, the originality of those ideas tended to increase (although this effect was somewhat inconsistent for the hygiene ideas).

As expected, cognitive stimulation during idea generation added to the effect of earlier knowledge activation. This again points to the relevance of extending my priming paradigm to interactive brainstorming groups. If knowledge activation before a brainstorming task enhances creative performance, and if being exposed to another group member's ideas can enhance this effect, interactive brainstorming groups may turn out to have an advantage over nominal brainstorming groups when it comes to deep exploration of subcategories.

## 3.4 Study 3.3: Effects of Priming in Interactive and Nominal Dyads

### *3.4.1 Introduction*

I conducted a study in which participants generated ideas in interactive or nominal dyads. In this study, I used my priming manipulation (nutrition or hygiene) in such a way as to create dyads that were homogeneous or heterogeneous with regard to priming.

In line with my earlier findings, I expected dyads that were primed homogeneously (i.e., dyads in which both participants were primed with one topic) to generate a higher percentage of ideas within that topic than dyads that were heterogeneously primed (i.e., where both participants were primed with two different topics) and control dyads (i.e., dyads that were not primed). I expected the originality of the generated ideas within the primed topics to follow the same pattern. Thus, I expected homogeneous dyads to generate the most original ideas within the primed topic, followed by heterogeneous dyads and control dyads.

Furthermore, because participants within an interactive dyad can influence each other's idea generation, I expected these effects to be significantly stronger for interactive dyads than for nominal dyads. Participants in nominal dyads are not exposed to the ideas of the other dyad member, and therefore cannot be influenced by the content of the other person's ideas. Thus, the effects of priming should simply be an addition of the independent effects for the two dyad members. In interactive dyads, however, both dyad members are exposed to the other person's ideas, and this should exacerbate the effect of priming: When both participants have been primed with the same topic, the priming effect should become even stronger than for a homogeneous nominal dyad. In effect, the ideas generated by another dyad member could re-activate the knowledge that was activated earlier through the priming manipulation. However, when the dyad members have been primed with different topics, this should decrease the effect of priming, because both dyad members are exposed to ideas from another topic than the topic they have been primed with. Thus, I expected the difference between homogeneous and heterogeneous dyads to be especially large for interactive dyads, and less so for nominal dyads.

In this study, I also included a selection task after idea generation. I was interested to see whether the poor selection performance found in Study 2.1 and Study 2.2 would be replicated using a different brainstorming topic. If my participants' earlier selection performance was due to their high personal involvement with the topic (presumably undergraduate Psychology students feel highly involved with education at the Department of Psychology), using another topic might allow them to take a more creative perspective and to select more original ideas. Alternatively, if my earlier results reflect something more fundamental in the way people select ideas after a brainstorming task, changing the brainstorming topic should not have any effect on the selection process, and participants' selection performance should again fail to exceed chance level.

### *3.4.2 Method*

#### *Participants*

Participants were 148 undergraduate students at Utrecht University (110 females and 38 males, mean age = 21.1 years), who received course credit.

#### *Design and Independent Variables*

This study employed a 2 (Type of Group: interactive or nominal) x 4 (Priming: homogeneous nutrition, homogeneous hygiene, heterogeneous or no priming) factorial design. Participants were randomly assigned to one of the conditions; all dyads were homogeneous with regard to gender.

*Type of Group.* Members of interactive dyads worked together during the idea generation and selection tasks. Members of nominal dyads worked simultaneously, but in separate rooms.

*Priming Questions.* The priming manipulation was identical to that used in the previous studies, except that this study was run as a paper-and-pencil task. Thus, all priming questions were filled out on a sheet of paper, instead of typed in on a computer. In homogeneous dyads, both dyad members were primed with questions about the same topic (i.e., nutrition or hygiene), while in heterogeneous dyads, one dyad member was primed with nutrition and the other was primed with hygiene.

*Procedure*

After entering the experimental room, participants individually read a general introduction to the experiment, in which it was explained that they would participate in an experiment regarding attitudes towards health. All participants (except participants in the no priming condition) then filled out the priming questions. Fifteen minutes were available for answering these questions. Next, participants received the instructions for the brainstorming task. As in the previous study, participants were told that they were to generate ideas in a brainstorming session, and that the goal of a brainstorming session is to come up with a large number of creative ideas. Participants were informed that the topic of the brainstorming session would be "what can people do to improve or maintain their health?" At this point, the four brainstorming rules were explained (quantity is wanted; freewheeling is welcomed; combine and improve ideas; criticism is ruled out). Participants were informed that twenty minutes would be available for idea generation. They were then assigned to a dyad and were brought to another room. Members of interactive dyads worked together in one room, members of nominal dyads worked separately in different rooms.

After the allotted twenty minutes had passed, the experimenter took away the lists of ideas and handed out the instructions for the selection task. In these instructions, it was explained that they were to select the four best ideas from the ones they had just generated. After participants had read these instructions, the ideas were handed back to the participants and they could make their selection. After the selection was made, participants were brought back to the first experimental room, where they filled in a short postexperimental questionnaire. Besides measuring participants' satisfaction with their performance, this questionnaire was intended to measure possible awareness of the link between my priming manipulation and the idea generation task; none of my participants gave any indication of such awareness. Next, participants were debriefed, thanked and paid.

*Measures and Dependent Variables*

*Global Productivity.* Productivity was defined as the total number of unique ideas generated by each dyad. For nominal dyads, this entailed adding the two sets of generated ideas from each dyad member, and removing the ideas that were identical in both sets. Two judges each scored half of the generated ideas for uniqueness. 380 ideas were scored by both judges, with agreement existing in 98.4% of the cases.

*Productivity Within Primed Topics.* I used the same procedure as in Study 3.1 to measure the percentages of ideas within the primed topics nutrition and hygiene. Thus, all ideas were first assigned a 'goal x means' code and were then classified into subcategories. A second rater coded 100 ideas in the 'goal x means system'. The raters assigned the same code in 92% of the cases. I deemed this to be sufficiently reliable and therefore used the first rater's codes.

*Idea Quality.* All ideas were scored on a 5-point scale by a trained rater for originality and feasibility. A second rater scored a subset of 250 ideas. With raters considered to be in agreement whenever their ratings differed by no more than one point, agreement existed in 98.7% of the cases for originality and in 96% of the cases for feasibility. Intraclass correlation coefficients were .77 for originality and .80 for feasibility. For each participant, I computed the mean originality and feasibility of the generated ideas as measures of global idea quality. I also computed the mean originality of the ideas generated within each of the primed topics (nutrition and hygiene).

With regard to the selected ideas, I followed a procedure similar to that used in Study 2.1. For interactive dyads, who selected four ideas together, I simply computed the average originality and feasibility of the four selected ideas. In contrast, members of nominal dyads selected four ideas individually, yielding eight selected ideas for each nominal dyad. I therefore computed the average originality and feasibility of the *four favorite* ideas of each nominal dyad, i.e., the ideas ranked 1 and 2 by each dyad member[5].

*Satisfaction.* Four items were used to measure participants' satisfaction with the number of generated ideas, the quality of their generated ideas, the quality of their selected ideas, and the way in which the selection was made. Participants indicated their response on a 9-point scale (1 = *not at all*, 9 = *very much*). All responses to the questionnaire items were averaged across dyad members to produce aggregate scores.

### *3.4.3 Results*

As in the previous studies, I employed contrast analyses to test specific hypotheses. Other dependent variables were analyzed with ANOVAs.

*Productivity*

*Global Productivity.* Table 3.4 presents the means and standard deviations for the dependent variables reported in Study 3.3. A 2 (Type of Group: interactive vs. nominal) x

---

[5] Similarly to Study 2.1, the average originality and feasibility of *all* selected ideas showed the same results; hence, this alternative measure is not reported.

4 (Priming: homogeneous nutrition, homogeneous hygiene, heterogeneous, no priming) ANOVA yielded a main effect of type of group: Nominal dyads generated more ideas ($M$ = 52.39, $SD$ = 20.41) than interactive dyads ($M$ = 28.11, $SD$ = 13.71; $F(1, 66) = 41.38$, $p < .001$, partial $\eta^2 = .39$). I also found a main effect of priming ($F(3, 66) = 4.24$, $p = .008$, partial $\eta^2 = .16$). I had no specific hypotheses with regard to the pattern of means, and performed post-hoc tests (Tukey's HSD) to test for differences between conditions. These revealed that dyads without priming generated more ideas ($M$ = 51.61, $SD$ = 26.94) than dyads with homogeneous nutrition priming ($M$ = 34.06, $SD$ = 17.76; $p = .010$) and dyads with homogeneous hygiene priming ($M$ = 35.89, $SD$ = 15.29; $p = .022$), and marginally significantly more than dyads with heterogeneous priming ($M$ = 38.42, $SD$ = 19.85; $p = .073$). Thus, while priming apparently reduced overall productivity, this effect was least strong for dyads who were primed heterogeneously. Finally, the interaction of priming and type of group was not significant ($F(3, 66) = 1.15$, $p = .34$), indicating that the effect of priming on productivity was the same for interactive and nominal dyads.

*Nutrition Ideas.* I first performed an ANOVA to determine whether the predicted interaction of priming and type of group was significant. A significant interaction would indicate that, as hypothesized, the effect of priming differed for interactive and nominal dyads. However, this was clearly not the case: The interaction of priming and type of group was not even remotely significant ($F < 1$). I therefore only tested my hypotheses with regard to the differences between the priming conditions.

Planned comparisons showed that, as expected, dyads with homogeneous nutrition priming generated a higher percentage of nutrition ideas ($M$ = 39.27, $SD$ = 20.61) than heterogeneous dyads ($M$ = 25.41, $SD$ = 15.25; $t(66) = 3.01$, $p = .002$, $r = .47$) and control dyads ($M$ = 17.96, $SD$ = 9.51; $t(66) = 4.70$, $p < .001$, $r = .73$), and that heterogeneous dyads generated a higher percentage of nutrition ideas than control dyads ($t(66) = 1.75$, $p = .043$, $r = .27$).

*Hygiene Ideas.* As with nutrition, an ANOVA with the percentage of hygiene ideas as dependent variable yielded no significant interaction of priming and type of group ($F < 1$). I therefore limited my analyses to the effects of my priming manipulation, collapsing conditions across type of group.

For the hygiene ideas, I anticipated a pattern of results similar to that of the nutrition ideas. Planned comparisons showed that, as expected, the homogeneous hygiene dyads generated a higher percentage of hygiene ideas ($M$ = 24.49, $SD$ = 18.66) than heterogeneous dyads ($M$ = 9.01, $SD$ = 13.55; $t(66) = 3.52$, $p < .001$, $r = .62$) and control dyads ($M$ = 7.09, $SD$ = 12.03; $t(66) = 3.96$, $p < .001$, $r = .70$). The difference between the heterogeneous and control dyads was not significant, however ($t < 1$).

Table 3.4 - Means and Standard Deviations for Study 3.3

| | *Nominal dyads* | | | |
|---|---|---|---|---|
| Measure | Nu / nu[a] | Hy / hy | Nu / hy | Control |
| Productivity | 45.44 (16.35) | 43.44 (13.97) | 52.11 (15.76) | 68.56 (25.95) |
| Perc. of nutrition ideas | 44.57 (25.26) | 19.57 (7.13) | 33.16 (18.25) | 22.92 (7.01) |
| Perc. of hygiene ideas | 1.09 (1.30) | 25.19 (17.53) | 13.49 (18.38) | 8.90 (16.44) |
| Originality of ideas [b] | 2.13 (0.41) | 1.97 (0.23) | 2.17 (0.30) | 1.96 (0.09) |
| Originality of nutrition ideas [b] | 1.91 (0.65) | 1.67 (0.39) | 2.00 (0.67) | 1.49 (0.25) |
| Originality of hygiene ideas [b c] | 2.00 (0.00) | 2.07 (0.08) | 2.29 (0.42) | 2.00 (0.00) |
| Satisfaction w. number of ideas[d] | 6.89 (1.17) | 5.89 (1.34) | 5.72 (1.18) | 6.2 (1.39) |
| Satisfaction w. quality of ideas[d] | 5.89 (0.93) | 5.89 (1.54) | 5.39 (1.54) | 5.89 (1.08) |
| Satisfaction w. selection process[d] | 7.28 (0.36) | 6.61 (0.49) | 6.67 (1.09) | 7.28 (0.87) |
| Satisfaction w. quality of selected ideas[d] | 6.39 (0.93) | 6.44 (1.16) | 6.33 (1.15) | 6.56 (1.47) |
| | *Interactive dyads* | | | |
| Measure | Nu / nu | Hy / hy | Nu / hy | Control |
| Productivity | 22.67 (10.55) | 29.10 (13.63) | 26.10 (14.53) | 34.67 (14.94) |
| Perc. of nutrition ideas | 33.98 (14.21) | 16.46 (9.58) | 18.43 (7.42) | 13.01 (9.37) |
| Perc. of hygiene ideas | 2.95 (4.48) | 23.87 (20.55) | 4.98 (5.37) | 5.28 (5.44) |
| Originality of ideas | 1.94 (0.54) | 2.14 (0.41) | 2.02 (0.38) | 1.94 (0.36) |
| Originality of nutrition ideas | 1.67 (0.73) | 1.55 (0.47) | 1.62 (0.64) | 1.33 (0.65) |
| Originality of hygiene ideas | 2.00 (0.00) | 2.15 (0.62) | 2.04 (0.10) | 2.00 (0.00) |
| Satisfaction w. number of ideas | 6.00 (1.25) | 6.25 (1.27) | 6.80 (1.38) | 6.94 (0.88) |
| Satisfaction w. quality of ideas | 6.11 (0.82) | 6.05 (1.52) | 7.05 (0.76) | 6.89 (0.89) |
| Satisfaction w. selection process | 7.39 (1.41) | 7.25 (1.11) | 7.70 (0.63) | 7.78 (0.57) |
| Satisfaction w. quality of selected ideas | 7.33 (0.43) | 6.75 (1.27) | 7.25 (1.14) | 7.67 (0.71) |

*Note.* $N$ = 74 dyads, except where otherwise indicated. Standard deviations are in parentheses.
[a] Nu / nu = Homogeneous nutrition; Hy / hy = homogeneous hygiene, Nu / hy = heterogeneous, control = no priming.
[b] Maximum value = 5.
[c] $N$ = 52.
[d] Maximum value = 9.

In summary, while my priming manipulation was successful, I did not find the expected interaction with type of group. That is, the effects of priming were not moderated by the presence or absence of actual interaction between the dyad members. Furthermore, the effects of my priming manipulation were most pronounced when both dyad members had been primed with the same topic. In dyads with heterogeneous priming, the effect of my manipulation seemed to disappear for hygiene ideas.

*Idea Quality*

*Global Idea Quality.* A mixed model ANOVA with the average originality and feasibility of the generated ideas as a within-subjects factor quality, and priming and type of group as between-subjects factors, yielded a main effect of quality ($F(1, 66) = 504.28$, $p < .001$, partial $\eta^2 = .88$), indicating that the generated ideas were, on average, lower in originality ($M = 2.04$, $SD = 0.36$) than in feasibility ($M = 4.22$, $SD = 0.48$). There were no further effects ($F$s $< 1$), indicating that the average originality and feasibility of the generated ideas were not affected by priming or by the presence or absence of dyadic interaction.

*Originality of Nutrition Ideas.* Planned comparisons for the originality of the nutrition ideas showed that, as expected, the nutrition ideas generated by homogeneous nutrition dyads were more original ($M = 1.79$, $SD = 0.68$) than the nutrition ideas generated by control dyads ($M = 1.41$, $SD = 0.48$; $t(66) = 1.99$, $p = .025$, $r = .67$). Furthermore, the nutrition ideas generated by heterogeneous dyads ($M = 1.79$, $SD = 0.66$) were also more original than those generated by the control dyads ($t(66) = 2.09$, $p = .02$, $r = .70$), while, unexpectedly, the difference between the homogeneous and heterogeneous dyads was not significant ($t < 1$).

*Originality of Hygiene Ideas.* Planned comparisons for the originality of the hygiene ideas showed that, contrary to expectations, the hygiene ideas generated by homogeneous hygiene dyads ($M = 2.12$, $SD = 0.47$) were no more original than those generated by control dyads ($M = 2.00$, $SD = 0.00$; $t < 1$). The hygiene ideas generated by heterogeneous dyads ($M = 2.18$, $SD = 0.34$) were marginally significantly more original than those generated by control dyads ($t(44) = 1.33$, $p = .09$). The difference between the homogeneous and heterogeneous dyads was not significant ($t < 1$).

*Correlations.* In line with the previous studies, the percentage of generated nutrition ideas was significantly correlated with the average originality of the nutrition ideas ($r = .55$, $p < .001$), and the percentage of hygiene ideas was significantly correlated with the originality of the hygiene ideas ($r = .41$, $p = .003$). The global originality of the generated ideas was not significantly related to overall productivity ($r = -.18$, $p = .12$), which again implies that it was not productivity as such that increased originality, but the relative productivity within a particular subcategory. Moreover, when I computed the correlation between global originality and overall productivity for nominal and interactive dyads separately, I actually found it to be significant and negative for nominal dyads ($r = -.36$, $p = .031$), and negative but nonsignificant for interactive dyads ($r = -.22$, $p = .19$).

*Selection Effectiveness*

In order to test the effectiveness of the selection process, I performed a mixed model ANOVA with the originality of the generated and selected ideas as two levels of a within-subjects factor originality, and priming and type of group as between-subjects factors. This analysis yielded only a significant effect of originality ($F(1, 66) = 55.39$, $p < .001$, partial $\eta^2 = .46$), indicating that, on average, the originality of the generated ideas ($M = 2.04$, $SD = 0.36$) was higher than the originality of the selected ideas ($M = 1.68$, $SD = 0.47$). This within-subjects effect was not qualified by an interaction with type of group, priming, or the interaction of these between-subjects factors (*F*s < 1), indicating that there were no differences between conditions with regard to the originality of the selected ideas.

A mixed model ANOVA with the average feasibility of the generated and selected ideas as two levels of a within-subjects factor feasibility, and priming and type of group as between-subjects factors, revealed a main effect of feasibility ($F(1, 66) = 8.23$, $p = .006$, partial $\eta^2 = .11$), indicating that the average feasibility of the generated ideas ($M = 4.22$, $SD = 0.48$) was lower than the average feasibility of the selected ideas ($M = 4.37$, $SD = 0.47$). As with originality, this within-subjects effect was not qualified by an interaction with type of group ($F(1, 66) = 1.70$, $p = .19$), priming ($F < 1$), or the interaction of these factors ($F < 1$).

In line with the previous studies, the average originality and the average feasibility of the generated ideas were strongly negatively correlated ($r = -.89$, $p < .001$), as were the average originality and feasibility of the selected ideas ($r = -.76$, $p < .001$). Furthermore, similar to my findings in Study 2.1, there was no correlation between productivity and the originality of the selected ideas ($r = -.10$, $p > .1$) and the feasibility of the selected ideas ($r = .11$, $p > .1$).

*Satisfaction*

I performed univariate 2 x 4 ANOVAs on the four items measuring participant satisfaction. Means and standard deviations are reported in Table 3.4. These analyses revealed that interactive dyads were more satisfied than nominal dyads with the quality of their generated ideas ($F(1, 66) = 7.69$, $p = .007$, partial $\eta^2 = .10$), with the selection process ($F(1, 66) = 7.68$, $p = .007$, partial $\eta^2 = .10$), and with the quality of the selected ideas ($F(1, 66) = 10.61$, $p = .002$, partial $\eta^2 = .14$). There was no difference between interactive and nominal dyads with regard to satisfaction with the number of generated ideas.

### *3.4.4 Discussion*

This study shows that my priming manipulation also worked within a group context, and most strongly when both group members were primed with the same topic. When the dyad members were primed with different topics, the effect of my manipulation on the relative production of ideas within the primed category largely disappeared. The fact that this was the case both in the interactive and the nominal dyads shows that this was not due to actual interaction (or a lack thereof) between the two group members. Instead, it is likely that the effect of priming was not very strong in magnitude and only became statistically significant when both group members contributed ideas within the primed topics to the dyad's production.

Unexpectedly, the originality of the nutrition ideas was as high for heterogeneous dyads as for homogeneous dyads. Furthermore, the originality of the hygiene ideas was highest for the heterogeneous dyads. Apparently, it was sufficient for one dyad member to be primed with a particular topic to increase the average originality of the dyad's ideas within that topic, even if the percentage of generated ideas within that topic was not necessarily higher. This again demonstrates that global measures of idea quantity and quality may obscure potentially important effects.

The finding that the presence or absence of actual dyadic interaction did not moderate the effects of my priming manipulation was, of course, unexpected. In Study 3.2, I found that exposure to ideas could influence the effect of priming, and I had expected to replicate this finding in an interactive setting. However, it should be noted that my stimulation manipulation in Study 3.2 was rather strong: For each idea that participants generated, a stimulation idea was presented. In retrospect, this may be something of an overestimation with regard to the cognitive stimulation and interference occurring in actual interactive dyads: It is not usually the case that every single idea generated by a dyad member is immediately followed by another idea from the other dyad member. Furthermore, as I noted above, the effects of group interaction are more complex; several social influences may have counteracted the effects of my priming manipulation.

With regard to idea selection, my results from Studies 2.1 and 2.2 were not exactly replicated: The quality of the selected ideas was not quite the same as the average quality of the generated ideas. The average feasibility of the selected ideas was higher than that of the generated ideas, but this does not necessarily mean that participants performed better than chance, since the average originality of the selected ideas was *lower* than that of the generated ideas. However, there was no difference between interactive and nominal groups with regard to the quality of the generated and selected ideas, or with regard to the effectiveness of the selection process. Furthermore, there was no correlation between productivity and the quality of the selected ideas. This is in line with my earlier findings.

## 3.5 General Discussion

In a brainstorming session, quantity is assumed to lead to quality, and research has confirmed that this relationship indeed exists. However, little empirical work has been done that sheds light on the question *why* quantity leads to quality. The chance view, which states that any generated idea has an equal chance of being a good, creative idea, and that the number of good ideas should therefore increase linearly with the total number of generated ideas, is supported by previous research. For example, correlations between productivity and the number of good ideas are usually quite high. However, in line with earlier suggestions by other researchers, I propose an alternative approach which states that creative idea generation depends on deep exploration of relevant domain knowledge, and that generating more ideas within a particular subcategory of the overall problem should be associated with a higher originality of those ideas. The studies presented in this chapter aimed to test this approach.

The degree to which people undertake deep exploration of a particular knowledge domain depends on the relative accessibility of their knowledge within that domain. Priming-type manipulations are an effective way to activate knowledge and thereby affect knowledge accessibility, and I therefore used such manipulations in these studies. Specifically, I had participants answer open-ended questions about subtopics (nutrition, hygiene or sports) of the ensuing brainstorming topic (health).

In Study 3.1, participants who had been primed with nutrition, hygiene or sports generated higher percentages of ideas within the primed categories than participants who had not been primed or who had been primed with an irrelevant topic. Furthermore, participants who had been primed with nutrition or sports also generated more original ideas within the primed topics. In contrast, global productivity and global originality were not affected by priming. Moreover, while global productivity and global originality were unrelated, the correlations between participants' percentages of nutrition and sports ideas and the originality of those ideas were significant. Thus, a higher relative productivity within a subcategory was associated with a higher average originality of the ideas within that subcategory (and with a higher number of original ideas within that subcategory).

In Study 3.2, I added cognitive stimulation to the paradigm. Participants were primed before brainstorming, and also received stimulation ideas during the brainstorming task (except participants in the control condition). I found that priming and stimulation both caused participants to generate a higher percentage of ideas within the activated topics, and most strongly so when priming and stimulation were homogeneous (i.e., when they concerned the same topic). Furthermore, for the ideas within the subcategory nutrition, the average originality of these ideas showed exactly the same pattern of means as the percentage of ideas within this subcategory. For the hygiene ideas, this effect was not as pronounced, but participants with hygiene priming and stimulation generated the highest percentage of hygiene ideas, and these also had the highest average originality.

In Study 3.3, I extended my priming manipulation to an interactive brainstorming situation. Participants in this study worked in interactive or nominal dyads, and were primed with the same topic (i.e., homogeneously), different topics (heterogeneously), or were not primed. As expected, homogeneous dyads generated the highest percentage of ideas within the primed subcategory, followed sequentially by heterogeneous dyads and dyads without priming. With regard to nutrition, both homogeneous and heterogeneous dyads generated more original ideas than dyads without priming. Heterogeneous dyads also generated the most original ideas about hygiene, while homogeneous dyads did not differ from the dyads without priming. Contrary to expectations, I found no effects of interaction between the dyad members. Thus, the stimulation effects found in Study 3.2 were not replicated with interacting dyads.

Several things can be concluded from these experiments. First, activating participants' knowledge with a priming manipulation can significantly affect creative performance. Earlier studies have found that priming can affect creativity via attentional processes and the activation of creativity-relevant stereotypes (Fitzsimons et al., 2005, in Chartrand, 2005; Friedman et al., 2003); my studies show that priming can also enhance creativity through the activation of domain knowledge.

Second, while previous research has suggested that productivity is not related to average idea quality (which in turn implies that every single idea has an equal chance of being a good idea), my experiments show that this relationship in fact does exist, but *within* subcategories. Clearly, global measures of quantity and quality give an incomplete picture of participants' creative performance, and thereby of the creative process. The chance approach to the quantity-quality relationship appears to be an adequate description on a global level, but when idea generation is studied from a cognitive perspective, things are more complex. In accordance with the reasoning of Finke, Ward, and Smith (1992), deep exploration of subcategories of a brainstorming problem can significantly increase the originality of the ideas generated within those subcategories.

Third, exposure to other ideas can add to the effects of knowledge activation, although I did not replicate this effect in an actual interactive setting. Earlier research has shown that exposure to ideas can affect the breadth and depth of idea generation (Nijstad et al., 2002). My studies show that the distinction between deep and broad brainstorming is not only relevant for the diversity of the generated ideas, but also for their originality.

*Limitations and Future Research*

*Causality in the Quantity-Quality Relation.* My results demonstrate that the quantity and quality of generated ideas are associated in more ways than has previously been assumed. However, the traditional view that idea quantity somehow *causes* idea quality implies mediation (with brainstorming leading to quantity, and quantity in turn leading to quality). I have found no evidence for such mediation in my data. The Baron and Kenny (1986) method of mediation analysis is highly conservative (e.g., Mackinnon, Lockwood, Hoffman, West, & Sheets, 2002) and the lack of support for a mediational interpretation might therefore be due to low statistical power. However, I propose that there may be a conceptual explanation, namely that the relation between quantity and quality is *not* causal, but correlational. If, for example, both quantity and quality are seen as outcome measures that reflect different dimensions of performance, it is hard to see how one could literally cause the other. Instead of assuming that quantity leads to quality, it would be more in line with my findings to state that both quantity and quality are enhanced by the same underlying processes. Thus, in the studies reported in this chapter, it appears that deeper processing of a subcategory caused higher productivity within that subcategory *and* caused the ideas generated within that subcategory to be more original.

*Feasibility of the Generated Ideas.* Idea quality is usually taken to consist of two components: originality and feasibility. Because my theoretical argument concerned only originality, I have not reported analyses for the feasibility of the generated ideas within subcategories. Furthermore, originality is often perceived to be strongly negatively correlated to feasibility (as indeed was the case in these studies), which might make such analyses redundant. Nevertheless, from an applied point of view, feasibility probably is as important as originality. I therefore performed post-hoc analyses with the feasibility of the generated ideas, and found that feasibility followed more or less the opposite pattern of originality: When participants generated more original ideas within a subcategory, these ideas generally were less feasible. Thus, the practical importance of my results might be questioned: What good is increasing the originality of ideas if it only happens at the cost of feasibility?

However, I think that this is not as problematic as it seems. First, the brainstorming procedure itself is specifically aimed at generating original and unusual ideas, without initial regard for practical feasibility. My results fit this goal very well. Second, adapting an unusual idea to fit practical constraints is probably easier, and more sensible, than adapting a conventional solution to be more original. Third, original ideas are not only important as solutions in their own right, but also because they may point to novel interpretations of a problem; these novel interpretations may themselves form a starting point for the generation of more feasible ideas. The creative process is a cyclical, rather than a linear process, and I speculate that originality may be a more promising starting point for new cycles than feasibility.

*Group Brainstorming.* My results suggest that exposure to another person's ideas can significantly affect the creative exploration of one's domain knowledge. However, I mainly found these results in a cognitive stimulation paradigm, and not in actual interactive dyads. It is possible that my priming manipulation was not strong enough to actually cause the expected stimulation effects in a face-to-face idea generation task. It has been shown that participants in brainstorming groups tend to spend their time rather inefficiently, and make many remarks that do not have direct bearing on the idea generation process (e.g., Putman & Paulus, 2003; also see Dugosh et al., 2000). Such distractions may weaken the effect of knowledge activation, and thereby the possible mutual stimulation or interference resulting from that stimulation.

*Choice of Topics.* The possibility of dividing a brainstorming topic into subcategories raises the question of *how* this division is to be made, and which topic and subcategories are to be used. This choice may significantly affect the results found. For example, subcategories must allow for extensive exploration of available knowledge. In my studies, participants generated very few ideas in the subcategory hygiene, severely limiting the statistical power of my analyses for idea quality. Furthermore, although I chose to use a more or less 'means-based' division of my brainstorming topic (nutrition, hygiene and sports can all be seen as means to the general end of health improvement), alternative divisions are of course possible (for example, using sub-goals, or taking the perspectives of different involved parties).

Moreover, the "health" problem itself may not be very conducive to creative idea generation, because it is such a broad and general problem, and because it lends itself so well to the retrieval of highly conventional instances from memory (such as "eat more fruit and vegetables," and "quit smoking"). It should be noted that, while participants in Studies 2.1 and 2.2. did not tend to select very creative ideas, participants in Study 3.3 actually selected ideas that were significantly *less* original (albeit significantly more feasible) than their generated ideas. Thus, changing the topic from education to health did make some difference for selection performance, but it cannot really be said that the selection became more creative. In general, given the wide variety of problems and tasks used in creativity research, it is surprising that no systematic study of problem properties and their effects on creative performance has yet been undertaken. This gap in the literature was noted 30 years ago by Stein (1975), and still exists today. A systematic exploration of problem properties and their effects on creative performance would be a promising and important line of future research.

*Implications*

*Deep Exploration and Fixation Effects.* As noted before, creativity is often associated with flexibility of thought, as opposed to rigidity and fixation. For example, the well-known experimental studies on *set* or *einstellung* effects (e.g., Luchins, 1942) show that people have a hard time deviating from an established problem-solving strategy, which can hinder their problem-solving efforts. Similarly, Jansson and Smith (1991) show that exposure to examples can hinder subsequent creative performance in a design task. In contrast, my results suggest that a certain degree of fixation may actually be beneficial to creative behavior, although the term *focus* would be more appropriate.

These views are not necessarily contradictory. Fixation as described in such studies as cited above, involves the continued use of specific elements of examples, or the continued use of previously adopted strategies. The problem in such cases is not that people search too deeply within a particular subcategory, but that they do not search deeply enough. As in the studies by Ward (1994) and Ward, Patterson, Sifonis et al. (2002), these "fixated" participants rely on solutions that are highly accessible and that seem obvious at the moment, and thus search their relevant domain knowledge only superficially. In my studies, I did not prime participants with specific ideas or solutions, but instead increased the general activation of (parts of) their relevant domain knowledge. Thus, while my participants probably still tended to rely on accessible knowledge items, this accessibility was not limited to a particular element or strategy, and therefore did not lead to fixation effects in the usual sense. However, it should be noted that in Study 3.3, dyads with homogeneous priming showed the lowest productivity. It is likely that knowledge activation can carry creative performance only so far. Although enhanced accessibility of domain knowledge makes it easier to keep generating ideas within a subcategory, this effect is not unlimited. At a certain point, the cognitive effort involved in generating new, creative ideas within a subcategory becomes too high. Moreover, when participants mainly generate ideas about a particular subcategory, domain knowledge in other subcategories may become inhibited, so that they find it hard to switch spontaneously to another subtopic and keep generating new ideas.

*Domain Knowledge and Creativity.* The results reported in this paper may be of interest in the debate regarding the relation between expertise and creativity. Despite the fact that domain knowledge is essential for true creativity to occur, this relation has often been a matter of contention. In a review chapter on knowledge and creativity, Weisberg (1999) distinguishes between the *tension view*, which states that domain knowledge and expertise cause people to rely in habits and common strategies, and thus hinder creative performance, and the *foundation view*, according to which extensive domain knowledge is necessary (but not sufficient) for high-quality creative work to occur.

The most well-known results in support of the tension view come from the research on *set* effects cited earlier. More recently, Wiley (1998) found that expertise was negatively related to performance on a remote associates task. Alternatively, Weisberg (1995) argued that notable creators such as Picasso and Edison strongly relied on the work of their predecessors and contemporaries, and that these influences were very important for their creative achievements. In line with this argument, Vincent, Decker and Mumford (2002) found that participants' expertise was positively related to various measures of creative performance.

In general, the detrimental effects on creativity do not seem to stem from domain knowledge per se, but from the way in which this knowledge is applied to the problem at hand; or, as Amabile (1996) put it: "while it is possible to have 'too many algorithms,' it is not possible to have too much knowledge" (p. 87). According to Amabile, if domain knowledge causes people to rely on well-known solutions or strategies, this is due to a lack of 'creativity-relevant skills.' Such skills can help people to break mental sets and scripts, to suspend judgment, etcetera. My results suggest that it is possible to induce a more creative use of available domain knowledge with a relatively small-scale intervention.

In conclusion, the results of my studies suggest that deep exploration of domain knowledge certainly has its merits, even though people normally do not tend to engage in such deep exploration. While the chance approach may adequately describe the relationship between global quantity and quality as it normally occurs, there clearly is more to idea generation than such global patterns. Deeper exploration of ideas within subcategories may not only be useful for participants in brainstorming studies, but for brainstorming researchers themselves as well.

CHAPTER 4

# Improving Selection Effectiveness [1]

## 4.1 Introduction

In the previous chapters, it has become clear that the availability of creative ideas is by no means sufficient to lead to the selection of creative ideas. In Study 2.1, I found that the ideas selected by interactive and nominal groups had the same average originality and feasibility as the ideas they had previously generated. In study 2.2, I found that participants who were provided with idea sets from Study 2.1 performed equally poorly, regardless of whether they were provided with explicit selection criteria. Finally, in Study 3.3, I found that interactive and nominal dyads selected ideas that were more feasible, but less original than the ideas they had previously generated. These results are also in line with other data on idea selection (e.g., Putman & Paulus, 2003).

Clearly, the potential for creative idea selection can easily be undone by an ineffective selection process. This raises the question which factors contributed to the ineffective selection performance found in earlier studies, and how selection effectiveness can be improved. This is the central topic of this chapter. I argue that an effective selection process rests on two factors. Firstly, every available idea must be thoroughly considered for selection. If some ideas are ignored, there is a chance that the best ideas are overlooked, which would lead to a suboptimal selection outcome. Secondly, the available ideas must be evaluated against specific and explicit selection criteria. If it is not clear what a "good" selection would be, it is impossible to optimize one's selection performance. The poor selection performance found in earlier studies may have been due to a) a lack of specific selection criteria, b) participants' failing to thoroughly consider and evaluate all available ideas, or c) a combination of these factors.

### *4.1.1 Selection Criteria*

A striking feature of most idea selection studies conducted so far (with the exception of Study 2.2; see below), is that participants in these studies were simply

---

[1] This chapter is based on Rietzschel, Nijstad, and Stroebe (2005b).

instructed to select the "best" ideas, or their "favorite" ideas (Faure, 2004; Putman & Paulus, 2003; Chapters 2 and 3 in this dissertation). However, idea quality can, theoretically, be judged on many dimensions. While the brainstorming instructions explicitly encourage the generation of creative and unusual ideas, participants in these studies were not informed that their performance would be rated specifically for originality and feasibility. It would be unfair to judge participants' performance as poor if it is not known what selection criteria they used, or if participants were not given the opportunity to adapt their performance to the relevant criteria.

Previous research has shown that people are able to calibrate their creative performance to specific instructions or criteria. For example, Harrington (1975) gave participants a divergent thinking test, in which they had to generate uses for familiar objects. Half of the participants were instructed to generate six *creative* uses (uses that were both unusual and worthwhile) whereas the other participants were simply instructed to generate six different uses. Participants with creativity instructions generated more creative uses, and fewer uncreative uses, than participants in the other condition.

Similarly, Shalley (1991) conducted a study in which participants had to generate solutions for problems in a fictitious steel manufacturing company. Participants were assigned productivity goals and/or creativity goals in a factorial design. Results showed that, on the whole, creativity goals enhanced the creativity of the proposed solutions (creativity, operationalized as the degree to which the solution was novel and appropriate, was rated on a 7-point scale), whereas productivity goals enhanced the quantity of the solutions.

Although these studies show that people are able to adapt their creative performance to specific instructions or goals, the tasks employed in these studies asked only for the *generation* of solutions or uses. It is not obvious that these results will generalize to the *selection* of generated alternatives. In Study 2.2 in this dissertation, I found that the ideas selected by participants with selection criteria were no more original or feasible than those selected by participants without selection criteria. However, this study did not include an idea generation stage; participants were presented with idea sets that had been generated by other participants in a previous experiment. The question thus remains whether people who first generate ideas and then have to select the best ideas from their own production, will make a better selection if provided with explicit selection criteria than participants who are simply instructed to select their "best" ideas.

### *4.1.2 Consideration of Available Alternatives*

Even given the availability of high-quality ideas and the presence of clear selection criteria, idea selection will be ineffective if the best ideas are not considered for selection. Ironically, the foundation of such neglect may be laid in the course of the brainstorming task. Although evaluation is not allowed during brainstorming, people find it difficult to refrain from evaluative thought and to engage in free speculation (Rickards, 1975), and it is likely that some of the proposed ideas will be viewed more favorably than others. The process of selecting the best ideas may then be focused on selecting from a set of favorite ideas, rather than from among all available ideas. Informal inspection of the video recordings made during Study 2.1 showed exactly such a pattern: Groups would often leaf quickly through the available ideas, until one or more of the group members would say something like "we should select this one," or "I thought that was a very good idea," after which the group would discuss whether the idea should be selected.

In general, people who have to make a choice from a large number of alternatives make use of some simplifying strategy, such as elimination by aspects (Tversky, 1972) or a satisficing strategy (Simon, 1955), in order to reduce the cognitive complexity of the decision (see Hastie & Dawes, 2001, for an overview). Even if their final goal is to select one single best option, people often engage in prechoice screening of the available alternatives (e.g., Beach, 1993): After participants have formed a consideration set of interesting or promising options, more cognitive resources are available for careful consideration of these remaining options (also see Parks & Cowlin, 1995). Brainstormers who want to select the best ideas from their own production will almost by definition have to choose from a large number of alternatives, and some simplifying strategy is therefore likely to be used. Because the ideas have been generated by the decision makers themselves, they are already familiar with the available alternatives and their relative attractiveness (unlike the usual situation in decision making experiments). In such a case, prechoice screening need not imply that all the alternatives are considered. Instead, participants are apt to process their generated ideas rather superficially, considering only those ideas that they remember from the idea generation stage as very good or appealing ideas, and neglecting (potentially good) ideas that were not immediately marked as favorites. Similarly, people who have to make a selection from a large set of pre-generated alternatives (as in Study 2.2) may be somewhat overwhelmed by the complexity of the task, and refrain from deep processing of all available ideas. Instead, they may focus on a salient attribute of these ideas (e.g., Karau & Kelly, 1992) to make a preselection and decrease the complexity of the selection task. This can undo the potentially beneficial effect of explicit selection criteria. If this is the case, selection criteria will only improve selection performance if participants are simultaneously induced to pay more attention to all available alternatives. This could be attained by having participants make a more careful preselection.

## 4.2 Study 4.1: Selection Criteria and Selection Procedure

### *4.2.1 Introduction*

In order to address the issue of selection criteria and systematic processing, I conducted a study in which participants first generated ideas, and then selected the best idea from their own production. I manipulated the presence of specific selection criteria and participants' preselection from among the available ideas.

*Selection Criteria*

The presence of selection criteria was manipulated by instructing half of the participants to select an idea that was both original and feasible. I used these two quality dimensions because, as explained in Chapter 1, quality in creativity research is usually defined as a combination of originality and feasibility. Hence, effective idea selection necessarily involves combining these dimensions. The other participants did not receive specific selection criteria, but were instructed to select "the best" idea.

*Preselection Instructions*

I further gave participants specific instructions for their preselection. When making a preselection from the available alternatives, people can use one of two basic strategies: They can either preselect the options from which they will make their final choice (positive), or eliminate the options which are unacceptable for further consideration (negative). Although these two strategies would seem to be logically equivalent, research shows that people do not treat them as such. For example, Shafir (1993) found that people treat available choice options differently, depending on whether they are instructed to *choose* an option, or to *reject* an option. Following up on this finding, Yaniv and Schul (1997) found that participants who were instructed to use a negative (i.e., elimination, or exclusion) preselection strategy formed a larger consideration set than participants who used a positive (i.e., inclusion) strategy. Furthermore, participants with elimination instructions more often chose the correct answer, presumably because a larger consideration set was more likely to contain the correct answer. This was confirmed in a study by Heller, Levin and Goransson (2002), who found that participants with exclusion instructions formed a larger consideration set, and made better choices, than participants with inclusion instructions, and that the latter effect disappeared when they controlled for the size of the consideration set.

If the poor selection performance found in earlier studies is due to participants' prematurely discarding good ideas, this problem might be avoided by giving participants exclusion instructions. Explicit instructions to form a consideration set would presumably lead participants to consider each individual idea, thus increasing the chance that available good ideas actually get selected. Furthermore, the larger consideration set formed under exclusion instructions as opposed to inclusion instructions is more likely to contain the best ideas, simply because it contains more ideas; this would increase the chance that the best ideas get selected (but see Levin, Huneke & Jasper, 2000, for an alternative view). Hence, I had my participants use an inclusion strategy (i.e., mark the ideas that were good enough for further consideration, prior to making the final selection) or an exclusion strategy (i.e., cross out the ideas that were *not* good enough for further selection, prior to making the final selection); I also employed a control condition where participants used neither an inclusion strategy nor an exclusion strategy.

I expected that participants with exclusion instructions would form a larger consideration set than participants with inclusion instructions. I further expected that the presence of selection criteria would lead to a more effective selection process, and that this in turn would lead to the selection of better ideas. However, I only expected these effects on selection to occur in combination with exclusion instructions.

### *4.2.2 Method*

#### *Participants*

Fifty-five undergraduate Psychology students (45 women, 10 men, mean age = 20.1 years) participated in this experiment. Participants received course credit or a reward of 7 Euros (about 9 US Dollars).

#### *Independent Variables*

The experiment had a 3 (Selection Procedure: inclusion, exclusion, or direct selection) x 2 (Instructions: criteria or no criteria) factorial design. Participants were assigned randomly to one of the experimental conditions. All participants worked individually throughout the whole experiment.

#### *Procedure*

The overall procedure consisted of the following tasks: idea generation, preselection, final selection, and exit questionnaire. The preselection task was performed only by participants in the inclusion and exclusion conditions; participants in the direct selection condition made their final selection immediately after the idea generation task.

Upon their arrival in the laboratory, participants received and read the first instructions (which were identical for all conditions). In these instructions it was explained that they were about to participate in a brainstorming session, and that they would generate ideas about possible improvements in the education at the Department of Psychology. The instructions then went on to explain that brainstorming is a technique to generate creative ideas and that the specific goal of a brainstorming session is to generate as many ideas as possible, without evaluating or criticizing these ideas. At this point, the four brainstorming rules (quantity is wanted; freewheeling is welcomed; combine and improve ideas; criticism is ruled out) were explained. Participants were advised to keep these rules in mind during idea generation; also, they were instructed to be concise and to the point in the phrasing of their ideas. They then generated ideas for 15 minutes, writing down their ideas on A4-sized sheets of paper. After this task, participants in the inclusion condition and the exclusion condition received the instructions for the preselection task, whereas participants in the direct selection condition received the instructions for the final selection task.

In the preselection task, participants were asked to read through the ideas they had just generated and to make a preselection from these ideas. Participants in the inclusion condition were instructed to mark those ideas that they thought were good enough for further consideration or elaboration; participants in the exclusion condition were instructed to cross out those ideas that they thought were *not* good enough for further consideration or elaboration. Half of the participants (the criteria condition) were told that "a good idea is both *original* (innovative and unusual) and *practically feasible*" and were instructed to use these criteria in making their preselection. The other participants did not receive specific criteria upon which to base their selection (the no criteria condition); these participants were simply instructed to mark [cross out] those ideas that were "good enough [not good enough] for further consideration." Participants were free to select or cross out as few or many ideas as they liked. No time limit was set for this task.

The next task consisted of making the final selection. Participants in the inclusion condition were instructed to select the best idea from those ideas that they had marked in the preselection task; participants in the exclusion condition were instructed to select the best idea from those ideas that remained after the preselection task; participants in the direct selection condition were instructed to select the best idea from the ideas they had generated in the idea generation task. In order to make sure that participants in the inclusion and exclusion conditions actually used the ideas selected or retained in the preselection task (instead of changing their preselection during the final selection task), participants received a red pen to make their final selection, whereas the preselection was made with a green pen. Participants in the criteria condition were again told that "a good idea is both *original* (innovative and unusual) and *practically feasible*" and were instructed to select the idea that best satisfied these criteria; the other participants were simply instructed to select "the best idea." No time limit was set for this task.

After making the final selection, all participants completed the postexperimental questionnaire. After completing this questionnaire, participants were debriefed, thanked and paid.

*Measures and Dependent Variables*

*Productivity.* Productivity was defined as the number of unique ideas generated by each individual.

*Consideration Set.* In the inclusion condition, the consideration set was defined as the number of ideas marked in the preselection task. In the exclusion condition, the consideration set was defined as the number of ideas retained (i.e., not crossed out) in the preselection task. Because participants in the direct selection condition went immediately to the final selection, no consideration set was computed for these participants.

*Idea Quality.* A trained rater who was blind to conditions coded all ideas for originality and feasibility. A second rater coded a random subset of 250 ideas. With raters considered to be in agreement whenever the ratings differed by no more than one point (Diehl & Stroebe, 1987), agreement existed in 97.6% of the cases for originality and in 99.6% of the cases for feasibility. As a more conservative measure of interrater agreement, I also computed the intraclass correlations, using a two-way random model and consistency definition (e.g., McGraw & Wong, 1996; Shrout & Fleiss, 1979); these were .74 for originality and .67 for feasibility. These values are considered "good" to "excellent" according to criteria specified by Cicchetti and Sparrow (1981). Hence, I concluded that the coding by my trained rater was reliable, and used these scores in my analyses.

I then computed two originality measures for each participant: the average originality of the generated ideas, and the originality of the idea selected in the final selection task. The same two measures were computed with the feasibility scores.

*Questionnaire Items.* All items in the questionnaire were rated on 5-point scales (1 = *not at all*, 5 = *very much*). Four items concerned selection criteria, asking participants to indicate whether they were expected to use originality (resp. feasibility) as a selection criterion (these two items served as direct manipulation checks), and whether they had found it important to use originality (resp. feasibility) as a selection criterion. Two other items measured selection satisfaction: One item asked participants whether they were satisfied with the way in which they made their selection, and one item asked participants whether they felt certain that they had selected the best idea.

### *4.2.3 Results*

*Manipulation Checks*

I analyzed the two manipulation checks with a 3 (Selection Procedure: inclusion, exclusion, or control) x 2 (Instructions: criteria or no criteria) ANOVA (see Table 4.1 for means and standard deviations). As expected, participants in the criteria condition more strongly felt that they were expected to use originality as a criterion ($M$ = 4.27, $SD$ = 0.72) than did participants in the no criteria condition ($M$ = 2.66, $SD$ = 1.14; $F(1, 49) = 40.00$, $p < .001$, partial $\eta^2 = .45$)[2], and participants in the criteria condition more strongly felt that they were expected to use feasibility as a criterion ($M$ = 4.12, $SD$ = 1.11) than participants in the no criteria condition ($M$ = 3.14, $SD$ = 1.25; $F(1, 49) = 9.07$; $p = .004$, partial $\eta^2 = .16$).

*Selection Criteria*

A 3 x 2 ANOVA showed that participants in the criteria condition found it more important to select original ideas ($M$ = 3.31, $SD$ = 0.88) than did participants in the no criteria condition ($M$ = 2.72, $SD$ = 0.92; $F(1, 49) = 5.13$; $p = .028$, partial $\eta^2 = .09$). In contrast, participants in the criteria condition found it just as important to select feasible ideas ($M$ = 4.08, $SD$ = 0.74) as did the participants in the no criteria condition ($M$ = 4.10, $SD$ = 0.77; $F < 1$).

I conclude that my criteria manipulation was successful. It should be noted, though, that participants in all conditions found it equally important to select feasible ideas, regardless of whether they were explicitly instructed to do so.

*Productivity*

Because my experimental manipulations were introduced after the idea generation task, I expected no productivity differences between conditions. As expected, a 3 (Selection Procedure) x 2 (Instructions) ANOVA yielded no significant effects (all $p$s > .1).

*Consideration Set*

Because participants in the direct selection condition did not form a consideration set, I performed a 2 (Selection Procedure: inclusion vs. exclusion) x 2 (Instructions)

---

[2] This ANOVA also revealed a marginal effect of selection procedure ($F(2, 49) = 2.93$, $p = .06$, partial $\eta^2 = .11$); however, post-hoc analyses showed that the differences between the inclusion, exclusion and control conditions were nonsignificant ($p > .1$).

ANOVA with consideration set as the dependent variable, omitting the direct selection condition. This analysis yielded the expected main effect of selection procedure: Participants in the exclusion condition formed a larger consideration set ($M$ = 10.00, $SD$ = 6.09) than participants in the inclusion condition ($M$ = 5.85, $SD$ = 2.87; $F(1, 35) = 7.25$, $p = .011$, partial $\eta^2 = .17$).

Table 4.1 - Means and Standard Deviations for Consideration Set, Idea Quality and Satisfaction in Study 4.1

| | Direct selection | | Inclusion | | Exclusion | |
|---|---|---|---|---|---|---|
| Measure | No criteria | Criteria | No criteria | Criteria | No criteria | Criteria |
| Ideas in consideration set | — | — | 4.90 (2.33) | 6.80 (3.16) | 10.90 (4.20) | 9.00 (7.84) |
| Originality of generated ideas[a] | 2.03 (0.46) | 2.14 (0.28) | 2.21 (0.47) | 2.19 (0.42) | 2.18 (0.45) | 2.08 (0.26) |
| Originality of selected ideas[a] | 2.11 (0.60) | 2.29 (1.11) | 2.10 (0.57) | 2.80 (0.79) | 2.20 (0.79) | 2.22 (0.67) |
| Feasibility of generated ideas[a] | 3.07 (0.48) | 3.06 (0.32) | 3.04 (0.29) | 3.02 (0.31) | 3.08 (0.42) | 3.24 (0.44) |
| Feasibility of selected ideas[a] | 3.22 (0.67) | 3.14 (1.07) | 3.20 (0.63) | 3.40 (0.52) | 3.20 (0.79) | 3.00 (1.00) |
| Satisf. with selection process[a] | 4.44 (0.53) | 4.00 (0.82) | 4.20 (0.42) | 3.90 (0.99) | 4.60 (0.52) | 4.33 (0.71) |
| Certainty of selection quality[a] | 3.56 (1.24) | 4.14 (0.69) | 3.40 (0.97) | 3.70 (1.16) | 4.00 (1.16) | 3.89 (0.93) |
| Was expected to select original ideas[a] | 3.00 (1.32) | 4.43 (0.53) | 2.10 (0.57) | 4.00 (0.82) | 2.90 (1.29) | 4.44 (0.73) |
| Was expected to select feasible ideas[a] | 2.78 (1.09) | 4.00 (1.00) | 3.10 (1.29) | 4.00 (1.25) | 3.50 (1.35) | 4.33 (1.12) |
| Important to select original ideas[a] | 3.11 (0.78) | 3.14 (0.69) | 2.30 (0.67) | 3.40 (0.97) | 2.80 (1.14) | 3.33 (1.00) |
| Important to select feasible ideas[a] | 3.78 (0.83) | 3.86 (0.38) | 4.40 (0.52) | 4.30 (0.67) | 4.10 (0.88) | 4.00 (1.00) |

*Note.* $N$ = 55. Standard deviations are in parentheses.
[a] Maximum value = 5.

*Idea Quality*

Means and standard deviations for idea quality are presented in Table 4.1. I performed a 3 (Selection Procedure) x 2 (Instructions) repeated measures ANOVA with the average originality of the generated ideas and the originality of the selected idea as the within-subjects factor originality. This analysis revealed no significant main or interaction effects (all $p$s > .1). Thus, there were no differences between conditions with regard to the originality of the generated ideas and the originality of the selected ideas. Furthermore, in none of the conditions was the originality of the selected ideas significantly different from the average originality of the generated ideas.

I ran a similar ANOVA with the average feasibility of the generated ideas and the feasibility of the selected idea as the within-subjects factor feasibility; again, no significant main or interaction effects were found (all *F*s $< 1$). As with originality, there were no differences between conditions with regard to the feasibility of the generated ideas and the feasibility of the selected ideas. Moreover, the feasibility of the selected ideas was not different from the average feasibility of the generated ideas; this was the case across conditions.

Thus, contrary to hypotheses, neither selection criteria nor selection procedure had any effect on participants' selection performance. In all conditions, the selection process was quite ineffective and not significantly better than chance. Correlational analysis further showed that productivity was not significantly related to the originality of the selected ideas ($r = -.18$, $p > .1$) or the feasibility of the selected ideas ($r = .08$ $p > .1$), indicating that participants with a higher productivity were not more likely to select high-quality ideas.

*Satisfaction*

I analyzed the two items that measured selection satisfaction with univariate 2 x 3 ANOVAs (see Table 4.1 for means and standard deviations). Participants in the criteria condition were less satisfied with the way in which they had made their selection ($M = 4.08$, $SD = 0.84$) than participants in the no criteria condition ($M = 4.41$, $SD = 0.50$), although this difference was only marginally significant ($F(1, 49) = 3.26$, $p = .077$, partial $\eta^2 = .06$). Participants in all conditions were equally certain that they had selected the best idea ($M = 3.75$, $SD = 1.02$; *F*s $< 1$).

### *4.2.4 Discussion*

This study replicated my earlier results (Rietzschel, Nijstad, & Stroebe, in press; see Chapter 2) with regard to idea selection: Participants selected ideas that were not significantly better than the average quality of their generated ideas. This was the case even when participants were explicitly instructed to take originality and feasibility into account, and also when participants were given exclusion instructions, which I had expected to lead to a better selection through the formation of a larger consideration set. Participants with selection criteria found it significantly more important to take originality into account while making their selection. However, participants with and without criteria found it equally important to select feasible ideas, which suggests that feasibility is a criterion that people tend to use spontaneously. Furthermore, participants who were given explicit selection criteria were slightly less satisfied with the way in which they made their selection, but all participants were equally certain that they had selected the best idea.

*Disagreement or Contradiction*

Although the results with regard to selection performance and the (lack of) effect of explicit criteria replicate my earlier findings, they raise the question of why selection criteria had no effect. One possible explanation for this discrepancy is that my view on originality and feasibility may have differed from that of my participants. I explicitly instructed participants to use specific selection criteria, but as long as I do not know the extent to which participants' quality judgments concur with mine, I still cannot be sure exactly what criteria they used. It is possible that my participants viewed other ideas as original or feasible than I did. This hypothesis could be tested by asking participants for explicit originality and feasibility ratings of the available ideas.

A second possible reason why selection criteria did not improve performance lies in the two-dimensional nature of my criteria instructions. It may have been difficult for participants to take both originality and feasibility into account while making their selection. For example, Manske and Davis (1968) provided participants with instructions to generate original and/or practical uses for a common object, and found that participants who were provided with both these criteria performed no better than a control group without criteria. Manske and Davis attribute this to the strong negative correlation between the average originality and practicality of the proposed uses ($r$ = -.80), which made it more or less impossible for their participants to satisfy both criteria at the same time. The same problem may have existed for my participants.

To test this explanation, I computed the correlation between the average originality and feasibility of the ideas generated in Study 4.1 and found it to be only marginally significant: $r$ = -.25, $p$ = .067. Thus, while this negative correlation may have contributed somewhat to my participants' low selection performance, it is unlikely that it can explain my results. Nevertheless, it is possible that participants *perceived* originality and feasibility to be negatively correlated dimensions, regardless of whether this correlation actually existed, and therefore found it difficult to reconcile these two seemingly incompatible selection criteria. This could also explain why participants with selection criteria were slightly less satisfied with the way in which they had made their selection: Participants may have perceived the selection criteria to be paradoxical and hence impossible to satisfy. If this explanation is correct, selection performance could be improved by using selection criteria that participants do *not* perceive to be contradictory.

*Preferences and Processing*

Although the exclusion instructions used in Study 4.1 did cause participants to form larger consideration sets, they had no effect on selection performance. Of course, it could be the case that, contrary to my hypothesis, all participants processed and considered the available alternatives quite thoroughly and hence had nothing to gain from using an exclusion strategy. However, it would be premature to draw that conclusion from this one study, especially in light of the fact that selection performance was so poor. It seems more plausible that simply changing the selection procedure was insufficient to actually change the way in which participants viewed and treated the available alternatives. For example, if participants had formed very strong preferences for particular ideas during idea generation, constructing a large consideration set would not necessarily alter these preferences. What would be needed is an intervention that induces participants to actively consider every single available idea with regard to relevant quality dimensions. If such an intervention does not improve selection performance, the previous research results on idea selection cannot be explained by participants' superficial processing of the available ideas.

## 4.3 Study 4.2: Idea Rating and Idea Selection

### *4.3.1 Introduction*

I conducted a study in which participants were presented with a pre-generated set of ideas, and were then asked to give quality ratings of these ideas, and to make a selection from among them. The primary aims of this study were to rule out the possibility that participants viewed other ideas as relatively original or feasible than I did, and to test my hypothesis that the lack of effect of criteria in Study 4.1 was caused by my dual selection criteria. The secondary aim of this study was to perform a new test of the hypothesis that deeper processing of the available ideas would improve selection performance.

*Creativity Instructions*

Even though researchers usually define creativity as a combination of originality (or novelty) and feasibility (or appropriateness), 'naive' participants may well regard originality as the single most important dimension of creative behavior and creative products, and as incompatible with feasibility or practicality (e.g., Runco & Charles, 1993). In Study 4.2, therefore, I instructed my participants to select *creative* ideas, without mentioning originality or feasibility. If most people regard originality as the most important dimension of creativity, instructions to select creative ideas would allow participants to focus on originality and hence enable them to improve their selection, at least with regard to that one dimension. Therefore, half of the participants in Study 4.2 were instructed to select the "most creative" idea, and half of the participants were instructed to select "the best" idea. I expected participants with creativity instructions to focus more strongly on originality while making their selection, to make a more effective selection, and to select more original ideas than participants without creativity instructions. Because the existing studies are ambiguous with regard to people's ability to effectively combine originality and feasibility considerations, I had no specific hypotheses with regard to feasibility.

*Quality Ratings and Processing*

Assuming that agreement exists between participants' and my own quality ratings, eliciting originality and feasibility judgments can also be a stronger way to induce deep processing of the available ideas. If so, participants who have made explicit quality judgments prior to making their selection should make a more effective selection than participants who have not made such a judgment. Specific quality judgments would make the originality and feasibility of every single idea salient and explicit, and therefore, presumably, would make participants more aware of the quality of the available ideas. Thus, half of my participants were instructed to rate all ideas for quality prior to making a selection, whereas the other participants immediately made a selection (and rated the ideas afterwards). I expected that the beneficial effect of creativity instructions would be strongest for those participants who had rated the ideas before making a selection.

### *4.3.2 Method*

*Participants*

Ninety undergraduate students initially participated in the experiment. However, two participants encountered software problems while performing the experimental task, and three participants did not follow instructions. Hence, the analyses reported below are based on 85 people (52 women, 33 men, mean age = 21,45 years). Participants received course credit or a reward of 7 Euros (about 9 US Dollars).

*Independent Variables and Materials*

The study was designed as a 2 (Instructions: creativity or default) x 2 (Task Order: select-first or rate-first) factorial design. Participants were assigned randomly to one of the experimental conditions. All participants worked individually throughout the experiment.

Because I was specifically interested in idea selection, participants in this study did not generate ideas, but were presented with a pre-generated set of 50 ideas. These ideas had been generated in previous experiments, where participants brainstormed about possible improvements in education at the department of Psychology. The ideas were selected to cover the entire range of originality and feasibility scores (from *not at all* to *very much*). The ideas were re-phrased where that was deemed necessary to understand the content of the idea, and spelling and typographical errors were corrected.

In order to make a comparison between my own and the participants' average ratings possible, the 50 ideas were rated for originality and feasibility by a trained rater and the author of this dissertation. Interrater reliability was high: The intraclass correlations (two-way random model, consistency definition) were .82 for originality and .83 for feasibility. The percentage of agreement was also high: With raters considered to be in agreement whenever the ratings did not differ by more than one point, the raters agreed in 98% of the cases for originality, and in 100% of the cases for feasibility. For each idea, I computed the average of the two raters' scores for both originality and feasibility.

*Procedure*

The experiment consisted of the following tasks: a rating task, a selection task, and a postexperimental questionnaire. Participants in the rate-first condition first performed the rating task and then the selection task; participants in the select-first condition first performed the selection task and then the rating task.

For the rating task, participants were seated in front of an Apple iMac computer. Ideas were presented individually, and the participant was asked to give his or her rating of those ideas. Ideas were to be rated on three dimensions: originality (i.e., the degree to which an idea was innovative and unusual), feasibility (i.e., the ease with which the idea could probably be realized) and desirability (i.e., the degree to which participants would like to see this idea implemented). I included desirability as a rating dimension (beside originality and feasibility), because I felt that participants would normally have a strong tendency to judge ideas in terms of their desirability. It seemed best, therefore, to anticipate on this tendency by explicitly including it in the experimental procedure. The program was set up in such a way that all ideas were rated first on one dimension, then on the next, and then on the last. Thus, for example, a participant would first rate all ideas (one by one) on originality, then he or she would rate all ideas on feasibility, and finally on desirability. Both the order of the dimensions (e.g., originality first, then feasibility, and then desirability) and the order of the ideas within each rating session were randomized.

For the selection task, participants received a printed list of 50 ideas (these were the same ideas that were also used in the rating task, but in the selection task all participants read the ideas in the same order). In the accompanying instructions, it was explained that these ideas came from a brainstorming session about possible improvements of education at the department of Psychology, and that their task would be to select the best four ideas from this set.[3] Participants were instructed to make their selection by assigning numbers to the four best ideas, in decreasing order of quality (thus, the best idea would receive number 1, the second-best idea would receive number 2, and so on). Half of the participants were instructed to "select the four most creative ideas" from this set (creativity condition), and half of the participants were simply instructed to "select the four best ideas" (default condition). Fifteen minutes were available for this selection task (this was enough for all participants).

After completing the postexperimental questionnaire, participants were debriefed, thanked and paid.

*Measures and Dependent Variables*

*Idea Quality.* I computed the average originality and feasibility (as perceived by the trained raters) of the four ideas selected by each participant.

*Questionnaire Items.* Five items assessed which selection criteria participants used. These items asked participants to indicate on a 5-point scale (1 = *not at all*, 5 = *very strongly*) whether they tried to select the most original idea, the most creative idea, the most feasible idea, the most desirable idea, and the best idea. One item assessed the degree to which participants were satisfied with the quality of the selected ideas.

### *4.3.3 Results*

*Agreement in Ratings*

In order to compare my participants' ratings of the 50 ideas in the idea set with those of the trained raters, I computed my participants' average originality and feasibility ratings of each idea contained in the set. Thus, I computed 50 average originality ratings (one for each single idea in the set) and 50 average feasibility ratings.

The intraclass correlation coefficient (average measure, two-way mixed, consistency definition) between the participants' average originality ratings and the trained raters' average originality ratings was .87, indicating that there was strong agreement between the trained raters and the untrained participants. Similarly, the intraclass coefficient between the participants' feasibility ratings and the trained raters' feasibility ratings was .82, also indicating strong agreement.

These results show that the trained raters and my participants viewed the same

[3] The reason for having participants select the four best ideas instead of one single best idea was that this would better approach the selection instructions as used in Study 2.1 and Study 3.3.

ideas as relatively original, and viewed the same ideas as relatively feasible. In other words, my earlier results probably cannot be explained by differences in opinion between the raters and my participants.

*Selection Criteria*

Table 4.2 presents the correlations between the different selection criteria used by my participants. Three results are particularly interesting. Firstly, as expected, my participants seemed to interpret "selecting the most creative ideas" as almost synonymous with "selecting the most original ideas" ($r = .839$, $p < .001$). Secondly, the degree to which participants tried to select "the best ideas" was highly correlated with the degree to which they tried to select the most feasible ideas ($r = .639$, $p < .001$) and the most desirable ideas ($r = .696$, $p < .001$). Thirdly, participants' selecting for originality was negatively correlated with their selecting for feasibility ($r = -.404$, $p < .001$).

Table 4.2 - Correlations Between Selection Criteria Used by Participants in Study 4.2

| | Most creative idea | Most original idea | Most feasible idea | Most desirable idea | Best idea |
|---|---|---|---|---|---|
| Most creative idea | — | | | | |
| Most original idea | .839** | — | | | |
| Most feasible idea | -.447** | -.404** | — | | |
| Most desirable idea | -.483** | -.476** | .759** | — | |
| Best idea | -.448** | -.403** | .639** | .696** | — |

*Note.* $N = 85$. ** = $p < .001$.

The means and standard deviations of all questionnaire items are presented in Table 4.3. A 2 (Instructions) x 2 (Task Order) ANOVA with participants' selecting for originality as the dependent variable yielded a main effect of instructions ($F(1, 81) = 78.09$, $p < .001$, partial $\eta^2 = .49$). As expected, participants with creativity instructions reported a stronger tendency to select for originality ($M = 4.05$, $SD = 0.87$) than participants with default instructions ($M = 2.25$, $SD = 1.01$). A main effect of instructions was also found for participants' selecting for creativity ($F(1, 81) = 70.22$, $p < .001$, partial $\eta^2 = .46$), indicating that participants with creativity instructions reported a stronger tendency to select creative ideas ($M = 4.41$, $SD = 0.84$) than participants with default instructions ($M = 2.61$, $SD = 1.15$). This main effect was qualified by an interaction between instructions and task order ($F(1, 81) = 4.82$, $p = .031$, partial $\eta^2 = .056$); simple effects analysis revealed that this effect was stronger in the select-first condition ($F(1, 81) = 56.63$, $p < .001$) than in the rate-first condition ($F(1, 81) = 18.88$, $p < .001$).

In contrast, participants with default instructions reported a stronger tendency to select feasible ideas ($M = 3.82$, $SD = 0.99$) than participants with creativity instructions ($M = 2.59$, $SD = 1.49$; $F(1, 81) = 20.85$, $p < .001$, partial $\eta^2 = .21$), and participants with default instructions reported a stronger tendency to select the "best" ideas ($M = 4.55$, $SD = 0.59$) than participants with creativity instructions ($M = 3.44$, $SD = 1.50$; $F(1, 81) = 20.46$, $p < .001$, partial $\eta^2 = .20$). In addition, participants with default instructions reported a stronger tendency to select desirable ideas (i.e., ideas that they would like to see implemented) ($M = 4.64$, $SD = 0.65$) than participants with creativity instructions ($M = 3.22$, $SD = 1.56$; $F(1, 81) = 30.01$, $p < .001$).

Table 4.3 - Means and Standard Deviations of Participants' Selection Criteria and Satisfaction in Study 4.2

| | Select first | | Rate first | |
|---|---|---|---|---|
| Measure | Default | Creativity | Default | Creativity |
| Most original idea | 2.27 (1.16) | 3.75 (0.85) | 2.23 (0.87) | 4.33 (0.79) |
| Most creative idea | 2.73 (1.08) | 4.05 (0.99) | 2.50 (1.23) | 4.76 (0.44) |
| Most feasible idea | 4.09 (0.81) | 2.70 (1.34) | 3.55 (1.10) | 2.48 (1.60) |
| Best idea | 4.73 (0.46) | 3.35 (1.42) | 4.36 (0.66) | 3.52 (1.60) |
| Most desirable idea | 4.59 (0.73) | 3.35 (1.46) | 4.68 (0.57) | 3.10 (1.67) |
| Satisf. with quality of selected idea | 4.32 (0.48) | 3.65 (1.14) | 4.23 (0.61) | 3.86 (1.06) |

*Note.* Maximum value = 5. $N = 85$. Standard deviations are in parentheses.

I conclude that my instructions manipulation was successful. Furthermore, these results confirm my hypothesis that participants perceived originality and feasibility to be contradictory selection criteria, and that creativity instructions caused participants to focus on originality. Moreover, participants who were instructed to "select the best ideas" preferred feasible, rather than creative, ideas. This is in line with the results from Study 4.1, where I found that participants tended to select for feasibility, regardless of whether they were explicitly instructed to do so.

*Idea Quality*

*Originality.* The means and standard deviations for the quality of the selected ideas are presented in Table 4.4. A 2 (Instructions) x 2 (Task Order) ANOVA with the mean originality of the selected ideas as the dependent variable yielded a main effect of instructions ($F(1, 81) = 41.31$, $p < .001$, partial $\eta^2 = .34$), indicating that, as expected, participants with creativity instructions selected more original ideas ($M = 2.78$, $SD = 0.73$) than did participants with default instructions ($M = 1.96$, $SD = 0.42$). There were no other significant main or interaction effects (*p*s > .1). Thus, contrary to my hypothesis, the effect of creativity instructions was not stronger for participants in the rate-first condition than for participants in the select-first condition.

Because the average originality of the idea set was constant (since each participant was presented with the same set of ideas), I could not use a repeated measures ANOVA (as I did in Study 4.1) to test for differences between the selection and the idea set. Instead, I performed separate t-tests for the participants with creativity instructions and the participants with default instructions, testing the originality of their selected ideas against the average originality of the idea set (= 2.18). These tests showed that participants with creativity instructions selected ideas that were more original than the idea set ($M = 2.78$, $SD = 0.73$; $t(40) = 5.27$, $p < .001$); in contrast, participants with default instructions selected ideas that were less original than the idea set ($M = 1.96$, $SD = 0.42$; $t(43) = -3.46$, $p = .001$).

These results confirm my hypotheses with regard to originality: Participants with creativity instructions selected more original ideas than participants with default instructions; they also selected ideas that were more original than the idea set, whereas the opposite was true for participants with default instructions.

Table 4.4 - Means and Standard Deviations for Idea Quality and Satisfaction in Study 4.2

| | Select first | | Rate first | |
|---|---|---|---|---|
| Measure | Default | Creativity | Default | Creativity |
| Originality of selected ideas | 1.92 (0.47) | 2.61 (0.69) | 2.00 (0.37) | 2.94 (0.74) |
| Feasibility of selected ideas | 2.87 (0.38) | 2.85 (0.34) | 3.13 (0.49) | 2.73 (0.45) |

*Note.* Maximum value = 5. $N = 85$. Standard deviations are in parentheses.

*Feasibility.* A 2 x 2 ANOVA with the mean feasibility of the selected ideas revealed a main effect of instructions ($F(1, 81) = 5.38$, $p = .023$, partial $\eta^2 = .06$); this main effect was qualified by an interaction between instructions and task order ($F(1, 81) = 4.18$, $p = .044$, partial $\eta^2 = .05$). Simple effects analysis revealed that, within the rate-first condition, participants with creativity instructions selected less feasible ideas ($M = 2.73$, $SD = 0.45$) than did participants with default instructions ($M = 3.13$, $SD = 0.49$; $F(1, 81) = 9.65$, $p = .003$); this difference was not found for the participants who first selected and then rated ideas ($F < 1$).

T-tests of the feasibility of the selected ideas against the average feasibility of the idea set (= 2.94) showed that, within the rate-first condition, participants with creativity instructions selected ideas that were less feasible than the idea set ($M = 2.73$, $SD = 0.45$; $t(20) = -2.16$, $p = .043$), whereas participants with default instructions selected ideas that were slightly more feasible than the idea set ($M = 3.13$, $SD = 0.49$), although this difference was only marginally statistically significant ($t(21) = 1.78$, $p = .089$). Participants within the select-first condition selected ideas that were neither more, nor less feasible than the idea set ($p$s > .1). Thus, creativity instructions did lead to the selection of ideas that were less feasible than the idea set, and less feasible than the ideas selected by participants with default instructions, but *only* within the group of participants who first rated the ideas for quality and then made their selection.

Finally, to test whether participants actually engaged in an originality-feasibility trade-off, I performed a 2 (Task Order) x 2 (Instructions) repeated measures ANOVA with the originality and the feasibility of the selected ideas as two levels of a within-subjects factor quality. A main effect of quality ($F(1, 81) = 33.17$, $p < .001$, partial $\eta^2 = .29$) was qualified by a two-way interaction of quality and instructions ($F(1, 81) = 31.72$, $p < .001$, partial $\eta^2 = .28$).[4] Simple effects analysis of the within-subjects effects revealed that participants with default instructions selected ideas that were more feasible ($M = 3.00$, $SD = 0.45$) than original ($M = 1.96$, $SD = 0.42$; $F(1, 81) = 67.28$, $p < .001$); participants with creativity instructions selected ideas that were as feasible ($M = 2.79$, $SD = 0.40$) as original ($M = 2.78$, $SD = 0.73$; $F < 1$).

*Mediation Model*

Apparently, creativity instructions increased participants' motivation to select for originality, and decreased their motivation to select for feasibility; this may in turn have led to a higher originality of the selected ideas. In order to test this mediation hypothesis, I performed mediation analyses in accordance with the procedure specified by Baron and Kenny (1986).

I first tested whether the effect of creativity instructions on the originality of the selection was mediated by the degree to which participants selected for originality. In a series of regression analyses, I found that instructions had a significant effect on the originality of the selection ($\beta = .575$, $t = 6.41$, $p < .001$; adjusted $R^2 = .32$) and that instructions had a significant effect on participants' selecting for originality ($\beta = .693$, $t = 8.77$, $p < .001$; adjusted $R^2 = .48$). When instructions and selecting for originality were

[4] This analysis also yielded a marginally significant three-way interaction of quality, instructions and task order ($F(1, 81) = 2.91$, $p = .092$), reflecting the fact that I found only a main effect of instructions on the originality of the selected ideas, and an interactive effect of task order and instructions on the feasibility of the selected ideas; these effects were discussed earlier.

simultaneously entered into a regression as predictors of originality of the selection, the effect of instructions decreased (β = .212, $t$ = 1.90, $p$ = .061), while selecting for originality had a significant effect (β = .524, $t$ = 4.72, $p < .001$; adjusted $R^2 = .46$). I performed a Sobel test (Mackinnon, Warsi, & Dwyer, 1995) to test whether the indirect path was significant, which it was: $Z = 4.15$, $p < .001$.

I then tested whether the effect of instructions on the originality of the selection was also mediated by a decreased tendency to select for feasibility. Instructions had a significant negative effect on participants' selecting for feasibility (β = -.448, $t$ = -4.565, $p < .001$, adjusted $R^2 = .19$). When instructions and selecting for feasibility were simultaneously entered into a regression as predictors of originality of the selection, the effect of instructions slightly decreased (β = .371, $t$ = 4.235, $p < .001$), while selecting for feasibility had a significant negative effect (β = -.455, $t$ = -5.190, $p < .001$, adjusted $R^2 = .48$). A Sobel test showed that the indirect path was significant: $Z = 3.43$, $p < .001$.

Finally, I entered instructions, selecting for originality and selecting for feasibility into a regression as predictors of the originality of the selection. The effect of instructions was no longer significant (β = .082, $t < 1$), while the effect of selecting for originality was significant (β = .452, $t$ = 4.585, $p < .001$), as was the (negative) effect of selecting for feasibility (β = -.402, $t$ = -5.062, $p < .001$, adjusted $R^2 = .59$). Thus, the effect of instructions on the originality of the selection was mediated by participants' tendency to use originality and feasibility as selection criteria; moreover, this mediation explained a sizeable proportion of the variance in the dependent variable.

*Satisfaction*

A 2 x 2 ANOVA revealed that participants with creativity instructions were less satisfied with the quality of the selected ideas ($M$ = 3.76, $SD$ = 1.09) than participants with default instructions ($M$ = 4.27, $SD$ = 0.54; $F(1, 81) = 5.71$, $p = .007$). Thus, participants who were instructed to select creative ideas were less satisfied about the quality of their selection. Again, this is in line with my earlier finding that participants tended to favor feasible ideas.

*4.3.4 Discussion*

The results of this study strongly suggest that the poor selection performance found in the criteria condition of Study 4.1 was indeed due to the dual nature of my selection criteria. When participants were simply instructed to select "creative" ideas, they were more motivated to use originality (as opposed to feasibility) as a selection criterion, and indeed selected more original ideas than participants who were instructed to select "the best" ideas. Furthermore, participants with creativity instructions made a more effective selection: They selected ideas that were more original than the idea set. This was not the case for participants with default instructions.

Participants showed a clear trade-off between the selection criteria originality and feasibility: The more they selected for originality, the less did they select for feasibility. This was not always reflected in their actual performance. When participants had not rated the ideas before they made their selection, creativity instructions only lead to an increase in originality of the selected ideas, without affecting the feasibility. However, when participants had rated the ideas prior to making their selection, creativity instructions increased the originality and decreased the feasibility of the selected ideas (as compared to default instructions). It is possible that giving quality ratings before selecting ideas actually strengthened participants' tendency to engage in an originality-feasibility trade-off: If participants perceive originality and feasibility to be strongly negatively correlated, and they are aware of the relative originality and feasibility of the available ideas, instructions to select creative ideas (which are interpreted as instructions to select *original* ideas) may lead them to seek out those ideas that are high on originality and low on feasibility. The perceived negative correlation between originality and feasibility would then become a self-fulfilling prophecy: Participants believe that they can only select original ideas at the cost of feasibility, and consequently, with the 'help' of their ratings, they do so. In any case, creative idea selection seems to benefit most from creativity or originality instructions without explicit quality ratings. Furthermore, the fact that quality ratings in no way enabled participants to make a better selection disconfirms my initial hypothesis that the earlier results on idea selection were largely due to superficial processing.

Interestingly, participants perceived the best ideas to be those ideas that are most practical or useful. Originality and creativity apparently were not dimensions that my participants would take into account spontaneously. In line with this view, participants with selection criteria were less satisfied with their selection. This contradicts my earlier hypothesis that participants were dissatisfied because of the perceived contradiction in the selection criteria. Instead, it seems that, whereas participants' natural tendency was to select for feasibility and desirability, creativity instructions (as well as the selection criteria used in Study 4.1) forced them to take originality into account. Because originality was perceived to be negatively related to feasibility, participants were required to select ideas that did not fit with their own preferences, and hence were less satisfied.

## 4.4 Study 4.3: Criteria and Problem Scope

### *4.4.1 Introduction*

The results from Study 4.1 and Study 4.2 show that people by default tend to select practical ideas, but are able to select better ideas when they receive creativity instructions. These results confirm my hypothesis that there is a lot to gain from enhancing selection effectiveness. Nevertheless, as I argued above, the quality of selected ideas also strongly depends on the quality of the available ideas. However, all participants in Study 4.1 received the same instructions for idea generation, and Study 4.2 did not incorporate an idea generation stage. I therefore conducted a third study that incorporated idea generation and idea selection, to test the combined effects of enhanced idea quality and creativity instructions. In this study, besides giving participants specific selection instructions, I manipulated the quality of the available ideas. I did this in two ways: by giving participants idea generation instructions in which either creativity or relevance was emphasized, and by varying the scope of the brainstorming topic.

*Creativity Versus Relevance*

As discussed above, previous research has shown that specific creativity instructions can improve the quality of generated ideas: When participants are instructed to be creative or original, they can adapt their performance to these instructions and generate more creative ideas or solutions. Furthermore, my previous results indicate that people tend to approach creativity and originality as synonymous. Hence, in the instructions for the idea generation task, for half of my participants I emphasized the importance of originality, and instructed them to generate ideas that were as original as possible. Similarly, for the idea selection task, these participants were instructed to select ideas that were as original as possible.

For the other participants, in contrast, I emphasized the *relevance* of the topic (improving education at the Department of Psychology), and instructed them to keep their experiences as students in mind while generating ideas. Similarly, for the idea selection task, they were instructed to base their selection on their personal experience. Research has shown that brainstorming is more difficult with a ‘relevant’ problem (i.e., a realistic problem that participants are actually interested in and care about) than with an irrelevant problem (Dillon, Graham & Aidells, 1972; Harari & Graham, 1975). Thus, when the task is more realistic and participants presumably experience more ‘real-world’ constraints, creativity appears to diminish. Explicitly emphasizing the relevance of the brainstorming topic, therefore, is likely to lead to less creative idea generation and selection than explicitly emphasizing the importance of originality.

I expected this manipulation to have three effects. Firstly, I expected participants with originality instructions to generate more original ideas than participants with relevance instructions. Secondly, because the originality of selected ideas is related to the average originality of the generated ideas, this higher availability of original ideas should also lead to the selection of more original ideas by participants with originality instructions. Thirdly, I expected participants with originality instructions to make a more *effective* selection: The originality of their selected ideas should be significantly higher than the average originality of their generated ideas. I did not expect such an effect for participants with relevance instructions. I had no specific expectations with regard to the average feasibility of the generated and selected ideas.

*Problem Scope*

Besides using creativity instructions, I also manipulated the scope of the brainstorming topic to increase my participants' brainstorming creativity. Again, if the quality of the selected ideas depends strongly on the quality of the available ideas, this manipulation of idea quality should also be reflected in the quality of the ideas that get selected.

Because my previous results show that, in the absence of an effective selection procedure, the average quality of the available ideas is the best predictor of the quality of the selected ideas, I was interested to see whether this relationship would hold when the average originality of the available ideas was explicitly manipulated, independently of whether participants were provided with criteria for idea quality. The fact that the quality of selected ideas thus far has consistently failed to be significantly higher than that of the generated ideas, suggests that it will: If people select ideas that are as good as their average ideas, the best ideas will be selected by those people who, on average, generate the best ideas.

Study 4.3 addressed this issue by using two different brainstorming topics, one of which was a sub-category of the other. Half of the participants in this study generated and selected ideas about a 'broad' problem (improving education at the Department of Psychology); the other participants generated and selected ideas about a 'narrow' problem (improving the lectures at the Department of Psychology). In the previous chapter, the deep exploration approach to the quantity-quality relationship in brainstorming was discussed. In three studies, I found that participants who generated relatively many ideas within a particular subcategory of the overall brainstorming problem also generated ideas of higher originality within these subcategories. In this study, therefore, I expected participants with a narrow problem scope to generate ideas of higher originality than participants with a broad problem scope, and I expected that this would also lead to the selection of more original ideas. However, I expected no effect of problem scope on the *effectiveness* of the selection process.

### *4.4.2 Method*

*Participants*

Initially, my sample consisted of 105 participants. However, three participants did not follow the experimental instructions and were excluded from the analyses. Hence, the analyses below are based on 102 undergraduate university students (76 women and 26 men, mean age = 21.1 years). Participants received course credit or a reward of 7 Euros (about 9 US Dollars).

*Independent Variables*

The study was designed as a 2 (Instructions: creativity or relevance) x 2 (Problem Scope: broad or narrow) factorial design. Participants were assigned randomly to one of the experimental conditions and worked individually throughout the whole experiment.

*Procedure*

The overall procedure consisted of the following tasks: idea generation, idea selection, and exit questionnaire.

Upon their arrival in the laboratory, all participants read a general introduction in which it was explained that they were about to participate in a brainstorming session, that the purpose of brainstorming is to come up with creative solutions or innovations, and that they would be required to first generate ideas, and then make a selection from these ideas. After this introduction, the brainstorming procedure was explained more fully: Participants read that brainstorming is a technique to generate creative ideas and that the specific goal of a brainstorming session is to generate as many original ideas as possible, without evaluating or criticizing these ideas. At this point, the four brainstorming rules were explained.

In the creativity conditions, the instructions then explained that the goal of the study was to find out how creative students could be in generating and selecting ideas, and that it would be their task to generate ideas that were as original (i.e., innovative and unusual) as possible. Alternatively, participants in the relevance conditions were told that the goal of the study was to find out how students generate and select ideas about a topic that they were personally involved in and had a personal interest in, and that it would be their task to generate as many ideas as possible.

Participants in the conditions with a broad problem scope were then told that they would be generating ideas about "possible improvements in the education at the Department of Psychology"; in the conditions with a narrow problem scope, the topic was "possible improvements in the lectures at the Department of Psychology". Participants in the creativity conditions were reminded that they were supposed to generate original ideas; participants in the relevance conditions were reminded to keep their experiences as a student in mind while generating ideas. Participants then generated ideas for 20 minutes, writing down their ideas on A4-sized sheets of paper.

After the brainstorming task, all participants received the instructions for the selection task. All participants were instructed to select the best idea from the ideas they had previously generated. Participants in the creativity conditions were instructed to base their selection on the originality of the ideas, whereas participants in the relevance conditions were instructed to base their selection on their experiences as students. A time limit of 15 minutes was set for this task.

After making their selection, all participants completed the postexperimental questionnaire, and were then debriefed, thanked and paid.

*Measures and Dependent Variables*

*Productivity.* Productivity was measured as the number of unique ideas generated by each individual.

*Idea Quality.* A trained rater who was blind to conditions rated all ideas for originality and feasibility. A second rater coded a random subset of 250 ideas on both dimensions. With raters considered to be in agreement whenever the ratings differed by no more than one point, agreement existed in 98% of the cases for both originality and feasibility. The intraclass correlation coefficients (two-way random model) were .73 for originality, and .63 for feasibility, which was satisfactory.

As in Study 4.1, I computed two originality and feasibility measures for each participant: the average originality (resp. feasibility) of the ideas generated by that participant, and the originality (resp. feasibility) of the idea selected by that participant. Furthermore, for each participant, I computed the number of generated high-quality ideas, with a high-quality idea defined as an idea that scored above the mean on both originality and feasibility.

*Questionnaire Items.* All items in the postexperimental questionnaire were rated on 5-point scales (1 = *not at all*, 5 = *very much*). Eight items were related to the idea generation task, and seven were related to the idea selection task.

Of the eight idea generation items, six were related to specific criteria; these asked participants how strongly they had tried to generate original ideas, creative ideas, feasible ideas, ideas that would offer advantages for all students, ideas that they personally would like to see implemented, and ideas that were solutions to problems that they themselves had encountered. Two items asked participants to indicate their satisfaction with the number of ideas they had generated and with the quality of the ideas that they had generated.

Of the seven idea selection items, six were related to the same criteria explained in the preceding paragraph, asking participants how strongly they had relied on these criteria in making their selection. One item asked participants how satisfied they were with the quality of the idea they had selected.

### *4.4.3 Results*

For clarity, I will first discuss the results with regard to idea generation and idea selection separately. I will then report the data on selection effectiveness.

*Idea Generation*

*Questionnaire Items.* Correlational analyses showed that participants' motivation to generate original ideas was positively correlated with their motivation to generate creative ideas ($r$ = .76, $p$ < .001) and negatively correlated with their motivation to generate feasible ideas, ($r$ = -.31, $p$ = .001). Furthermore, participants' motivation to generate feasible ideas was positively correlated with their motivation to generate ideas that would offer advantages for students ($r$ = .21, $p$ = .034), that they would like to see implemented ($r$ = .35, $p$ < .001), and that would be solutions to problems they had encountered ($r$ = .37, $p$ < .001).

Table 4.5 presents means and standard deviations for the questionnaire items. A 2 x 2 ANOVA showed that participants with a narrow brainstorming topic reported a slightly stronger tendency to generate creative ideas ($M$ = 4.08, $SD$ = 0.76) than participants with a broad brainstorming topic ($M$ = 3.79, $SD$ = 1.04), although this effect was only marginally significant ($F$(1, 98) = 2.78, $p$ = .10, partial $\eta^2$ = .03). Participants in all conditions found it equally important to generate original ideas ($M$ = 3.61, $SD$ = 1.04, $p$s > .1). Participants with relevance instructions reported a stronger tendency to generate ideas that would offer advantages for students ($M$ = 4.29, $SD$ = 0.76) than participants with originality instructions ($M$ = 3.75, $SD$ = 0.99; $F$(1, 98) = 9.21, $p$ = .003, partial $\eta^2$ = .09). Participants with relevance instructions also reported a slightly stronger tendency to generate ideas that they would like to see implemented ($M$ = 4.39, $SD$ = 0.76) than participants with originality instructions ($M$ = 4.06, $SD$ = 1.12; $F$(1, 98) = 2.98, $p$ = .087, partial $\eta^2$ = .03); they also reported a stronger tendency to generate ideas that were solutions to problems that they had encountered (*M* = 4.29, *SD* = 0.94) than participants with originality instructions (*M* = 3.75, *SD* = 1.18; $F(1, 98) = 6.61$, $p = .012$, partial $\eta^2$ = .06). The latter item also yielded a main effect of problem scope ($F(1, 98) = 6.49$, $p$ = .012, partial $\eta^2$ = .06): Participants with a broad topic reported a stronger tendency to generate solutions to experienced problems (*M* = 4.27, *SD* = 0.93) than participants with a narrow topic (*M* = 3.74, *SD* = 1.19).

Table 4.5 - Means and Standard Deviations for Criteria Used by Participants in Study 4.3

| | Broad problem | | Narrow problem | |
|---|---|---|---|---|
| Measure | Relevance | Originality | Relevance | Originality |
| Generated original ideas | 3.44 (1.12) | 3.48 (1.19) | 3.79 (0.88) | 3.73 (0.92) |
| Generated creative ideas | 3.84 (0.99) | 3.74 (1.09) | 4.13 (0.68) | 4.04 (0.66) |
| Generated feasible ideas | 3.72 (0.84) | 3.44 (1.19) | 3.71 (0.95) | 3.38 (1.29) |
| Generated ideas with advantages for students | 4.12 (0.88) | 3.85 (1.03) | 4.46 (0.59) | 3.65 (0.98) |
| Generated ideas they wanted to see implemented | 4.52 (0.71) | 4.15 (1.17) | 4.25 (0.79) | 3.96 (1.08) |
| Generated solutions to problems they had encountered | 4.52 (0.65) | 4.04 (1.09) | 4.04 (1.12) | 3.46 (1.21) |

*Note.* Maximum value = 5. *N* = 85. Standard deviations are in parentheses.

Thus, participants with relevance instructions clearly focused more strongly on generating ideas that would be feasible and useful. In contrast, participants with originality instructions reported a slightly stronger tendency to generate creative ideas. However, all participants found it equally important to generate original ideas; this was probably due to the fact that originality was emphasized in the brainstorming instructions that all participants received. Similarly to my findings in Study 4.2, participants' motivation to generate creative, original ideas was negatively related to their motivation to generate feasible ideas. In contrast, the motivation to generate feasible ideas was positively related to the motivation to generate useful or 'relevant' ideas.

*Productivity.* A 2 (Instructions) x 2 (Problem Scope) ANOVA yielded no significant main or interaction effects, indicating that neither originality instructions nor problem scope had an effect on the number of ideas that participants generated.

Table 4.6 - Means and Standard Deviations for Idea Quality in Study 4.3

| | Broad problem | | Narrow problem | |
|---|---|---|---|---|
| Measure | Relevance | Originality | Relevance | Originality |
| Originality of generated ideas | 2.15 (0.23) | 2.21 (0.21) | 2.23 (0.37) | 2.43 (0.43) |
| Originality of selected idea | 2.12 (0.83) | 2.52 (0.89) | 2.12 (0.85) | 2.77 (0.77) |
| Feasibility of generated ideas | 3.29 (0.34) | 3.25 (0.23) | 3.44 (0.33) | 3.29 (0.45) |
| Feasibility of selected idea | 3.24 (0.72) | 3.19 (0.83) | 3.33 (0.82) | 3.15 (0.78) |

*Note.* Maximum value = 5. $N$ = 85. Standard deviations are in parentheses.

*Idea Quality.* All means and standard deviations with regard to idea quality are presented in Table 4.6. Contrary to my findings in Study 4.1, the average originality and feasibility of the ideas generated in this study were significantly negatively correlated ($r = -.57$, $p < .001$). A 2 x 2 ANOVA with the originality of the generated ideas as dependent variable yielded a significant main effect of problem scope ($F(1, 98) = 5.56$, $p = .02$, partial $\eta^2 = .054$): As expected, participants who generated ideas about a narrow problem generated more original ideas ($M = 2.33$, $SD = 0.41$) than did participants who generated ideas about a broad problem ($M = 2.18$, $SD = 0.23$). This analysis also yielded a marginally significant main effect of instructions ($F(1, 98) = 3.81$, $p = .054$, partial $\eta^2 = .037$), indicating that, also in line with expectations, participants with originality instructions tended to generate more original ideas ($M = 2.31$, $SD = 0.35$) than participants with relevance instructions ($M = 2.19$, $SD = 0.32$).

An ANOVA with the feasibility of the generated ideas as the dependent variable yielded no significant main or interaction effects ($p$s > .1). Thus, neither problem scope nor instructions had any effect on the feasibility of the generated ideas.

An ANOVA with the number of high-quality ideas as the dependent variable yielded no significant main or interaction effects ($F$s < 1), which is in line with the result for productivity.

In line with my expectations, originality instructions and a narrow brainstorming topic both increased the average originality of the generated ideas. However, the effect of instructions was only marginally significant. As with the questionnaire items, this might be due to the fact that all participants received brainstorming instructions, and that these instructions always emphasize originality. Similarly to my findings in Study 4.2, manipulations that increased the average originality of ideas did not decrease their average feasibility, despite the negative correlation between these two dimensions. This again suggests that it is possible to improve the creative quality of ideas by focusing on one specific dimension.

Table 4.7 - Means and Standard Deviations for Criteria Used by Participants and Satisfaction in Study 4.3

| | Broad problem | | Narrow problem | |
|---|---|---|---|---|
| Measure | Relevance | Originality | Relevance | Originality |
| Selected original ideas | 2.32 (1.28) | 3.22 (1.37) | 2.46 (1.22) | 3.81 (1.06) |
| Selected creative ideas | 2.92 (1.04 ) | 3.33 (1.36) | 2.83 (1.09) | 3.92 (0.93) |
| Selected feasible ideas | 4.08 (0.91) | 3.29 (1.20) | 4.04 (0.91) | 3.42 (1.10) |
| Selected ideas that would offer advantages for students | 4.40 (0.71) | 4.15 (1.03) | 4.67 (0.65) | 3.58 (1.24) |
| Selected ideas that they would like to see implemented | 4.60 (0.76) | 4.48 (0.80) | 4.75 (0.53) | 4.38 (0.75) |
| Selected solutions to problems they had encountered | 4.36 (1.04) | 3.89 (1.42) | 4.33 (0.87) | 3.11 (1.21) |
| Satisfaction with number of ideas generated | 3.56 (0.92) | 3.00 (0.96) | 3.67 (0.64) | 3.27 (0.87) |
| Satisfaction with quality of ideas generated | 3.88 (0.78) | 3.11 (1.19) | 3.58 (0.78) | 3.19 (0.85) |
| Satisfaction with quality of selected idea | 4.44 (0.65) | 3.85 (0.99) | 4.21 (0.66) | 3.96 (0.77) |

*Note*. Maximum value = 5. $N$ = 85. Standard deviations are in parentheses.

*Idea Selection*

*Questionnaire Items*. Correlational analyses showed that participants' motivation to select original ideas was positively correlated with their motivation to select creative ideas ($r$ = .73, $p < .001$) and negatively correlated with their motivation to select feasible ideas, ($r$ = -.24, $p$ = .015). Furthermore, participants' motivation to select feasible ideas was positively correlated with their motivation to select ideas that would offer advantages for students ($r$ = .33, $p$ = .001), and marginally correlated with their motivation to select ideas that would be solutions to problems they had encountered ($r$ = .18, $p$ = .076).

Means and standard deviations for questionnaire items are reported in Table 4.7. ANOVAs showed that participants with originality instructions reported a stronger tendency to select original ideas ($M$ = 3.51, $SD$ = 1.25) than participants with relevance instructions ($M$ = 2.39, $SD$ = 1.24; $F(1, 98)$ = 21.06, $p < .001$, partial $\eta^2$ = .18) and they reported a stronger tendency to select creative ideas ($M$ = 3.62, $SD$ = 1.19) than participants with relevance instructions ($M$ = 2.88, $SD$ = 1.05; $F(1, 98)$ = 11.45, $p$ = .001, partial $\eta^2$ = .11). Conversely, participants with relevance instructions reported a stronger tendency to select feasible ideas ($M$ = 4.06, $SD$ = 0.89; $F(1, 98)$ = 11.48, $p$ = .001, partial $\eta^2$ = .11) and a marginally stronger tendency to select ideas that they would like to see implemented ($M$ = 4.67, $SD$ = 0.66; $F(1\ 98) = 2.84$, $p = .095$, partial $\eta^2 = .03$) than participants with originality instructions ($M$ = 3.36, $SD$ = 1.15; $M$ = 4.43, $SD$ = 0.77, respectively). The item measuring whether participants selected ideas that would offer advantages for all students yielded a significant main effect of instructions ($F(1, 98) = 13.19$, $p < .001$, partial $\eta^2 = .12$), that was qualified by an interaction of instructions and problem scope ($F(1, 98) = 5.15$, $p = .026$, partial $\eta^2 = .05$). Simple effects analysis revealed that, within the 'narrow topic' condition, participants with relevance instructions reported a stronger tendency to select ideas that would offer advantages for all students ($M$ = 4.67, $SD$ = 0.56) than participants with originality instructions ($M$ = 3.58, $SD$ = 1.24) $F(1, 98) = 17.07$, $p < .001$); this difference was not significant in the 'broad topic' condition. Thus, whereas participants with originality instructions focused more on selecting original and creative ideas, participants with relevance instructions tended to focus more on feasible and useful ideas.

*Idea Quality.* Means and standard deviations for idea quality are presented in Table 4.6. The average originality and feasibility of the selected ideas were significantly negatively correlated ($r = -.25$, $p = .012$), although this correlation was significantly smaller than for the generated ideas ($Z = 2.76$, $p = .003$). The originality of the selected ideas was not correlated with productivity ($r = -.08$, $p > .1$) or with the number of generated high-quality ideas ($r = .11$, $p = .29$), but the correlations between the feasibility of the selected ideas and productivity ($r = .22$, $p = .03$) and the number of generated high-quality ideas ($r = .24$, $p = .02$) were significant, indicating that participants who generated more ideas tended to select ideas that were higher in feasibility.

A 2 x 2 ANOVA with the originality of the selected ideas as the dependent variable yielded only a main effect of instructions ($F(1, 98) = 9.88$, $p = .002$, partial $\eta^2 = .09$): Participants with originality instructions selected more original ideas ($M$ = 2.64, $SD$ = 0.83) than participants with relevance instructions ($M$ = 2.12, $SD$ = 0.83); contrary to expectations, there was no effect of problem scope.

An ANOVA with the feasibility of the selected ideas as dependent variable yielded no significant effects ($F$s < 1): Neither instructions nor problem scope had any effect on the feasibility of the selected ideas.

Thus, whereas participants with creativity instructions selected more original ideas than participants with relevance instructions, problem scope had no effect on the originality of the selected ideas.

*Selection Effectiveness*

In order to test whether the originality of the selected ideas differed from the average originality of the generated ideas, I conducted a 2 x 2 x 2 mixed model ANOVA with the average originality of the generated ideas and the selected idea as the within-subjects factor originality. This analysis revealed a significant Originality x Instructions interaction ($F(1, 98) = 6.25$, $p = .014$, partial $\eta^2 = .06$). Simple effects analysis showed that participants with originality instructions selected ideas that were more original ($M = 2.64$, $SD = 0.83$) than their generated ideas ($M = 2.31$, $SD = 0.35$; $F(1, 98) = 8.91$, $p = .004$); this difference was not significant for participants with relevance instructions ($F < 1$).

A repeated measures ANOVA with the feasibility of the generated and selected ideas as the within-subjects factor feasibility yielded no significant effects (all *p*s > .1): The feasibility of the selected ideas was not different from that of the generated ideas, and this was the case in all conditions.

Thus, selection effectiveness was improved by giving participants originality instructions for the selection task: Participants with originality instructions selected ideas that were more original, and not less feasible, than their idea set, whereas participants with relevance instructions performed at chance level.

*Satisfaction*

The three items measuring participant satisfaction were analyzed with univariate 2 x 2 ANOVAs (means and standard deviations are presented in Table 4.7). For all three items, a main effect of instructions was found: Participants with relevance instructions were significantly more satisfied with the number of ideas they generated ($F(1, 98) = 7.87$, $p = .006$, partial $\eta^2 = .07$), the quality of the ideas they generated ($F(1, 98) = 10.09$, $p = .002$, partial $\eta^2 = .09$) and the quality of the idea they selected ($F(1, 98) = 7.19$, $p = .009$, partial $\eta^2 = .07$) than participants with creativity instructions.

Correlational analyses revealed an interesting pattern. Participants' satisfaction with the quality of the ideas they had generated was positively correlated with the degree to which they tried to generate ideas that they would like to see implemented ($r = .24$, $p = .015$) and ideas that would be solutions to problems they had encountered ($r = .27$, $p = .006$), and marginally correlated with the degree to which they tried to generate feasible ideas ($r = .18$, $p = .064$). Interestingly, participants' satisfaction with the number of ideas they generated was also positively correlated with the degree to which they tried to generate original ideas ($r = .34$, $p < .001$) and creative ideas ($r = .31$, $p = .001$). Similarly, participants' satisfaction with the quality of their generated ideas was marginally significantly correlated with the degree to which they tried to generate original ideas ($r = .17$, $p = .084$) and significantly correlated with the degree to which they tried to generate creative ideas ($r = .26$, $p = .009$).

Conversely, participants' satisfaction with the quality of their selected idea was negatively correlated with their motivation to select an original idea ($r = -.25$, $p = .011$)

and marginally negatively correlated with their motivation to select a creative idea ($r$ = -.17, $p$ = .098). In contrast, participants' satisfaction with the quality of their selected idea was positively correlated with their tendency to select a feasible idea ($r$ = .26, $p$ = .008), an idea that they would like to see implemented ($r$ = .34, $p$ = .001), an idea that would be a solution to problems they had encountered ($r$ = .38, $p$ < .001) and ideas that would offer advantages for students ($r$ = .21, $p$ = .036).

Thus, all three measures of satisfaction show the same pattern (in ANOVAs) as in Study 4.2 (i.e., participants with originality instructions were less satisfied with their performance). In line with these results, the correlations show that participants who strove for feasible, 'relevant' ideas were generally more satisfied with their production and their selection. Striving for originality was associated with lower satisfaction in idea selection, but not in idea generation.

### *4.4.4 Discussion*

The results of Study 4.3 replicate and further extend my findings from Study 4.2. Originality instructions enhanced the quality of generated and selected ideas, and improved the effectiveness of the selection process. Participants with originality instructions were able to make an effective selection from the ideas that they had generated, while participants with relevance instructions performed at chance level. Furthermore, although participants who generated ideas about a narrow topic also generated more original ideas, this had no effect on the quality of the selected ideas. Thus, improving the average quality of the available ideas is not sufficient to improve the quality of the selected ideas.

Similarly to my findings in Study 4.2, participants' self-reports with regard to their selection criteria show a trade-off between originality and feasibility: The tendency to select for originality was strongly negatively correlated with the tendency to select for feasibility. Again, this confirms that the low selection effectiveness found in Study 4.1 was due to the dual nature of the selection criteria. As in Study 4.2, instructions that caused participants to focus on originality did reduce their motivation to generate and select feasible ideas, but this was not reflected in their actual performance.

The results with regard to satisfaction also replicate my earlier results. Participants with originality instructions were less satisfied with all aspects of their performance (productivity, quality of production and quality of selection) than participants with relevance instructions. However, the pattern of correlations shows that, on the whole, focusing on originality and creativity during idea generation caused greater satisfaction with productivity and the quality of their generated ideas. This is possibly due to participants' perception of the task: The brainstorming instructions, which were given to all participants, encourage the generation of as many creative ideas as possible. Participants who actually tried do so may have felt that they met the task demands, and may therefore have been more satisfied with their performance. In contrast, during idea selection, focusing on originality and creativity was associated with lower satisfaction.

## 4.5 General Discussion

While nominal groups clearly generate more ideas than interactive groups, this advantage of nominal groups is not consistently reflected in their selection of creative ideas. Instead, the existing studies suggest that both interactive and nominal groups perform very poorly on idea selection. I therefore conducted three studies to gain insight into the selection process and the ways in which selection might be improved.

I argued that an effective selection is only possible if all available ideas are taken into consideration and evaluated with regard to a relevant quality dimension. I hypothesized that these two factors were responsible for the poor selection performance found in earlier studies. Studies 4.1 and 4.2 were designed to test these hypotheses.

In Study 4.1, I employed exclusion instructions (as compared with inclusion instructions and a control condition) to increase my participants' processing of the available ideas, and crossed this manipulation with the presence (vs. absence) of explicit selection criteria (originality and feasibility). Neither of these manipulations had an effect on the quality of the selected ideas. Furthermore, in none of the conditions was the quality of the selected ideas different from the average quality of the generated ideas.

I then hypothesized that the lack of effect of criteria in Study 4.1 was due to a perceived incompatibility of these dual criteria. In Study 4.2, therefore, half of the participants were simply instructed to select the most creative ideas. In addition, I asked my participants for explicit quality ratings of a pre-generated set of ideas, either before or after making a selection from these ideas. I found that creativity instructions improved selection performance (i.e., increased the originality of the selected ideas), but decreased satisfaction. Giving quality ratings before making a selection did not improve participants' performance. Thus, while consideration of the available alternatives arguably is essential for effective idea selection, this does not appear to have been a major bottleneck in the existing studies. Instead, the main cause of the poor selection performance found earlier seems to have been the lack of specific and non-contradictory selection criteria.

In Study 4.3 I addressed the question whether improving the average quality of the generated ideas would also lead to a higher quality of selected ideas. Two experimental manipulations (a narrow problem scope and specific originality instructions) did increase the average originality of participants' ideas, but this in itself did not lead to the selection of more original ideas. Instead, originality instructions for the selection stage did cause participants to select more original ideas, and ideas that were more original than their idea sets.

While the studies reported in this chapter do not answer all questions with regard to idea selection, several issues are very clear. The generation of creative ideas in a brainstorming session does not automatically lead to the selection of creative ideas. Moreover, improving the average quality of the generated ideas in itself has no effect on the quality of the selected ideas. However, the selection process can be improved by giving people instructions to focus on selecting original ideas. When left to their own devices, people appear to prefer feasible or useful ideas, and they perceive originality to be negatively correlated with feasibility. Therefore, instructing people to select ideas that are both original and feasible does not improve selection performance. Nevertheless, people are able to recognize their original ideas and, given the right instructions, to select them. Furthermore, selecting original ideas does not necessarily mean selecting less feasible ideas. Thus, despite the perceived (or real, as in Study 4.3) negative correlation between originality and feasibility, an effective combination of these dimensions is possible. However, if people do not use the "right" selection criteria, their inclination to select feasible ideas will undo the creative potential of their initial production. In addition, because people apparently prefer feasible, rather than creative, ideas, they may be less satisfied with their selection if they are instructed to focus on originality.

In retrospect, it is perhaps not very surprising that improving the average quality of the generated ideas did not improve the quality of the selected ideas. Recall that interactive groups in Study 2.1 did select ideas that differed marginally significantly from their average production, undoing the initial quality differences between nominal and interactive groups. This could reflect a regression to the mean: Even though people may generate ideas of different average quality under different circumstances, they will still select ideas of more or less equal quality, because their selection criteria and preferences have not changed. Thus, only an intervention that changes the criteria used during idea selection affects the quality of the generated ideas.

An obvious limitation of these studies is that I only used individual participants. The question whether interactive groups have a benefit to offer in idea selection remains unanswered. Although the existing research has failed to show a consistent group benefit in idea selection, research by Laughlin and colleagues, also mentioned in Chapter 2, has shown that groups can outperform individuals on decision tasks (e.g., Laughlin, Bonner, & Altermatt, 1998; Laughlin & Shippy, 1983; Laughlin, VanderStoep, & Hollingshead, 1991). These studies used intellective tasks with demonstrably correct options. The selection of creative ideas, in contrast, is a judgmental task with no single correct option (Laughlin, 1980). However, when the selection process is to be based upon a specific criterion, and if group members agree on the definition and operationalization of this criterion, the task becomes more intellective; this could confer a benefit upon groups as compared to individuals. Group members could also call attention to an original idea that other group members failed to consider, or they could give new information that renders a highly original idea more feasible, making it more likely that such an idea will be accepted by the group members. On the other hand, Kaplan and Miller (1987) found that a judgmental (as opposed to an intellective) task caused decision-making groups to rely more strongly on normative influence, and less on informational influence. This may detract group members' attention from the relative merits of each generated idea, leading to suboptimal idea selection. This is all speculative, and as stated above, the available data on idea selection do not show any advantage (nor a consistent disadvantage) for interactive groups. Nevertheless, idea selection in interactive groups remains an interesting and important avenue for further research, especially given the widespread use of brainstorming groups and in light of the continuing controversy regarding the effectiveness of interactive groups as compared with individuals and nominal groups.

When studied in isolation, a brainstorming session is effective when a large pool of creative ideas is generated. In context, however, a brainstorming session is effective when this pool of ideas results in the selection and possibly the implementation of one or more creative ideas. Furthermore, brainstorming as an idea generation tool is *useful* only when increasing the pool of available ideas also increases the likelihood that creative ideas are selected. Previous research has failed to demonstrate a consistent advantage of nominal over interactive groups with regard to idea selection. Thus, the most productive groups do not necessarily select the best ideas. To test whether these results were replicated in the three studies reported in this chapter, I computed the correlation coefficients between the number of ideas generated by my participants and the quality of their selected ideas. In Study 4.1, neither the originality nor the feasibility of the selected ideas was significantly correlated with productivity. In Study 4.3, I found an effect of originality instructions on selection performance, and therefore computed these correlations separately for participants with relevance instructions and participants with originality instructions. For participants with relevance instructions, only the feasibility of the selected ideas was (marginally) correlated with productivity ($r$ = .25, $p$ = .077); this correlation was not significant for the originality of the selected ideas. For participants with originality instructions, neither the originality nor the feasibility of the selected ideas were correlated with productivity.[5]

Thus, in line with earlier findings, a higher productivity was not associated with the selection of more creative solutions or ideas. In this sense, then, quantity does *not* breed quality; that is, increasing the availability of high-quality ideas seems to be irrelevant for the quality of selected ideas.

[5] Since Study 4.2 employed an identical idea set for all participants, there was no variance in productivity; hence, no correlation could be computed.

CHAPTER 5

# Summary and Conclusions

## 5.1 Introduction

Brainstorming is a well-known and highly popular method for generating creative ideas. The main principles behind brainstorming are the belief that quantity breeds quality, and the deferment of judgment. Adherence to the brainstorming rules, which are derived from these principles, should increase the number of ideas that are generated in a brainstorming session, and should therefore increase the number of available high-quality ideas, and the chance that a good idea gets selected for further development.

Brainstorming research has focused on productivity (i.e., the number of generated ideas), with special attention for the robust finding that interactive brainstorming groups generate fewer ideas than the same number of people working separately whose unique ideas are pooled (nominal brainstorming groups). As a consequence, little is known about the degree to which creative idea generation in a brainstorming session actually leads to the selection of creative ideas, and the question remains whether "more is better" when idea generation is studied in the context of subsequent idea selection. The studies reported in this dissertation addressed this issue. In Chapter 2, I reported a study in which I compared interactive and nominal groups on idea generation and selection, to test the hypothesis that the higher productivity of nominal groups would lead to their selecting better ideas. In Chapter 3, I reported three studies that looked into the relationship between quantity and quality, and the effects of deep exploration of domain knowledge on idea generation and selection. In Chapter 4, I addressed the issue of selection effectiveness, and factors that influence selection performance.

The general approach of these studies (although there were differences) was to ask participants to generate ideas about a particular topic, and then to ask them to select the best ideas from their generated ideas. I would then count the unique ideas, and compute the average originality (the degree to which ideas are unusual and innovative, as rated on a 5-point scale by a judge) and feasibility (the degree to which ideas are practical and the ease with which they could be implemented, also as rated by a judge) for all generated and selected ideas. These measures then allowed me to compare the quality of the selected ideas to the quality of the generated ideas, and to test the relationship between quantity and quality.

## 5.2 Summary of Main Findings

### *5.2.1 Idea Selection in Interactive and Nominal Groups*

If boosting productivity indeed increases the chance of selecting high-quality ideas, I should expect nominal brainstorming groups, who are the most productive, to select the best ideas. In Study 2.1, therefore, I compared interactive and nominal groups on a task that involved both idea generation and idea selection. Participants in this study (undergraduate Psychology students) generated and selected ideas about the question how education at the Department of Psychology could be improved. The separation between these two tasks was either emphasized by presenting idea generation and idea selection as two separate tasks, or de-emphasized by presenting them as two aspects of one task. I found that task separation affected productivity most strongly in interactive groups: When idea generation and selection were presented as one task, interactive groups generated fewer ideas per minute than when idea generation and selection were presented as two separate tasks. While nominal groups clearly generated more ideas than interactive groups (especially when idea generation and selection were presented as one task), and hence had more high-quality ideas at their disposal, this higher productivity did not lead to their selecting better ideas. In fact, although the ideas generated by nominal groups were, on average, more original and less feasible than those generated by interactive groups, nominal and interactive groups selected ideas of equal quality. The disappearance of the initial quality differences was due to the interactive groups' selecting ideas that were marginally more original, and marginally less feasible, than their generated ideas. On the whole, however, the selected ideas were not very creative, and selection performance was not significantly better than chance. Moreover, the quality of the selected ideas was not significantly related to the number of generated ideas, or to the number of available high-quality ideas.

Clearly, the higher productivity of nominal groups as compared to interactive groups does not necessarily lead to the selection of better ideas. The poor selection performance displayed by participants in this study was remarkable, and this raised the question whether participants would select more creative ideas if they were provided with explicit selection criteria. This hypothesis was tested in Study 2.2. Participants in this study were presented with randomly selected idea sets from Study 2.1, and were instructed to select the best ideas from this set. Half of the participants in this study were specifically instructed to use originality and feasibility as selection criteria, and the other participants were simply instructed to select “the best” ideas. However, these instructions did not affect participants’ selection performance: Participants with selection criteria did not select better ideas than participants without selection criteria, and selection performance again was not better than chance. This makes it unlikely that the poor selection performance found in Study 2.1 was due to a lack of explicit selection criteria.

The main conclusion to be drawn from these studies is that the benefits of nominal groups’ higher productivity (as compared to interactive groups) can easily be undone by an ineffective selection process: Productivity is not enough. Further, the average quality of selected ideas is unrelated to the availability of high-quality ideas, and selection criteria are not sufficient to improve selection effectiveness.

### *5.2.2 Knowledge Activation and Deep Exploration of Domain Knowledge*

Although Study 2.1 did not reveal any benefit of a high productivity when it came to idea selection, it remains indisputable that the quality of selected ideas is constrained by the quality of the available ideas. Previous research has shown that productivity is related to the availability of high-quality ideas, but little is known about the nature of this relationship. According to the chance or Darwinian view of creativity, each generated idea has an equal or near-equal chance of being a good idea (e.g., Simonton, 1999; also see the ensuing commentaries). Thus, although a higher productivity should lead to a higher number of generated high-quality ideas, the proportion of good ideas and the average quality of the generated ideas should be unaffected by productivity. An alternative view proposes that the most creative ideas are generated when people engage in deep exploration of their knowledge about a particular topic or problem, but that this is not something people tend to do spontaneously (e.g., Finke, Ward, & Smith, 1992). In contrast to the Darwinian view, the deep exploration view suggests that a higher productivity can be associated with the generation of ideas of higher originality. As people generate more ideas within a particular domain, they should go beyond the highly accessible ideas, which are also the most conventional (and hence unoriginal) ideas, and generate more original ideas.

In line with the Darwinian approach to idea generation, research has shown that the total number of generated ideas is linearly related to the number of good ideas contained therein, and that productivity generally is unrelated to the average quality of generated ideas. However, these studies have used only global measures of quantity and quality. In contrast, the studies reported in Chapter 3 addressed the question whether a high relative productivity *within a subcategory* of the overall brainstorming topic is associated with higher originality for those specific ideas. If so, this would argue against a strict Darwinian approach to the relationship between quantity and quality in idea generation.

In all three studies reported in Chapter 3, I used a priming-type manipulation to increase the activation level of participants' knowledge about specific subcategories of the brainstorming topic, and thereby induce deeper exploration of those subcategories. Specifically, the priming manipulation entailed a series of open-ended questions that participants had to answer; these questions concerned a subtopic of the brainstorming topic. The brainstorming topic concerned ways to improve or maintain one's health; the subcategories of this topic were nutrition, hygiene, and (in Study 3.1) sports.

In Study 3.1, priming increased both the relative productivity within the primed subcategories and the average originality of those ideas. Thus, participants who were primed with a particular subcategory of the brainstorming topic generated relatively more ideas about that subcategory, and these ideas (with the exception of the ideas within the subcategory hygiene) were also more original. Global productivity and idea quality, however, were not affected by priming.

The results of Study 3.1 show that deep exploration of domain knowledge, as measured by the relative productivity within a subcategory of the brainstorming topic, indeed is associated with higher originality; which argues against a strictly Darwinian view of creativity. I then hypothesized that being exposed to other participants' ideas would influence this deep exploration. Hence, in Study 3.2, besides employing the same priming manipulation, I provided participants with stimulation ideas during the brainstorming task. The priming effects from the earlier study were largely replicated, and these effects were particularly strong when priming and stimulation were homogeneous, i.e., when priming and stimulation concerned the same topic.

These results suggest that exposure to another person's ideas can indeed add to the effects of priming on productivity and originality within subcategories. The next question was whether I would find similar results in an actual interactive setting. Furthermore, I wanted to test whether the effects of knowledge activation on idea generation would carry over to idea selection. In Study 3.3, therefore, I had participants generate and select ideas in interactive or nominal dyads that were homogeneous or heterogeneous with regard to priming. In homogeneous dyads, both dyad members had been primed with the same topic; in heterogeneous dyads, the dyad members had been primed with different topics. Productivity within the subcategories nutrition and hygiene was highest for homogeneously primed dyads, followed by heterogeneously primed dyads. The nutrition ideas generated by the dyads with heterogeneous priming and the dyads with homogeneous nutrition priming were more original than those generated by dyads without priming, while the hygiene ideas generated by the heterogeneous dyads were more original than those generated by dyads with homogeneous hygiene priming and dyads without priming. Contrary to my expectations, the presence or absence of dyadic interaction did not make any difference for these priming effects. With regard to global productivity, I found that (besides the normal productivity loss in interactive dyads) homogeneously primed dyads were the least productive, while dyads without priming were the most productive. Again, this effect was not moderated by dyadic interaction. Further, I found no effects of type of group or priming on idea selection: In all conditions, participants selected ideas that were less original, and more feasible, than their generated ideas.

The results of Study 3.3, then, show that priming can also lead to deeper exploration of subcategories of the brainstorming problem in a dyadic context. However, I did not find evidence for stimulation or interference effects due to interaction, which is not in line with the results of Study 3.2. It is possible that the stimulation manipulation in Study 3.2 was stronger than the ‘real’ effects of dyadic interaction, and that other group processes counteracted the potential stimulation or interference in interactive dyads.

In sum, these studies support the notion that deeper exploration of available knowledge can increase the originality of the ideas that people generate. Furthermore, a priming manipulation, which increases the accessibility of specific domain knowledge, can induce such deeper exploration without changing the nature of the brainstorming task. However, the effects of deep exploration apparently do not carry over to idea selection.

### *5.2.3 Selection Effectiveness*

Thus far, my results showed that the generation of creative ideas was not sufficient to lead to the selection of creative ideas, and that selection performance was far from optimal. Therefore, Chapter 4 went more deeply into the issue of selection effectiveness. In this chapter, I argued that an effective selection entails that each available alternative be carefully evaluated against specific criteria. The poor selection performance found in my earlier studies may have been due to superficial processing, a lack of specific criteria, or a combination of these two factors. When making a selection from a large number of alternatives, people are likely to use a simplifying strategy to reduce the cognitive load of the task. This may cause them to process the available alternatives suboptimally. Even if explicit selection criteria are provided, as was the case in Study 2.2, it is possible that these will only improve the selection process if people are motivated to engage in more thorough consideration of the available ideas.

The goal of Study 4.1 was to test these hypotheses. Participants in this study first generated and then selected ideas (the topic was the "education" topic also employed in Chapter 2). During idea selection, participants were instructed to form a *consideration set* through an inclusion strategy (mark the best ideas and then make a selection from those ideas) or an exclusion strategy (cross out the worst ideas and then make a selection from the remaining ideas), or to make a direct selection. Based on previous research, I hypothesized that participants with exclusion instructions would form a relatively large consideration set; I further expected that this would lead them to consider more of the available ideas for selection, and this in turn was expected to lead to the selection of more creative ideas. Furthermore, half of the participants were provided with explicit selection criteria (originality and feasibility). Although I had found no effect of selection criteria in Study 2.2, I expected that selection criteria in combination with a more thorough consideration of the available ideas would enhance selection performance. However, neither of these manipulations affected selection performance: There were no differences between conditions with regard to the quality of the selected ideas, and participants in all conditions performed at chance level. Thus, my findings from Studies 2.1 and 2.2 were replicated, despite my manipulation of the selection procedure.

I reasoned that this result could have been caused by three factors. Firstly, instructions to take both originality and feasibility into account might have been perceived as contradictory, preventing a beneficial effect of selection criteria. Instead of forcing participants to use two specific criteria, it would be illuminating to know what quality dimensions they would take into account when they try to select creative ideas. Secondly, an exclusion strategy might not have been sufficient to cause more thorough processing of the available ideas. Instead, inducing participants to explicitly consider the quality of each single available idea might be sufficient to reach this goal. Thirdly, it was possible that participants viewed different ideas as original and feasible than I did. If so, I cannot say that they made an ineffective selection. Thus, I needed to asses the agreement between my raters' quality scores and participants' perceptions of quality. To test these hypotheses, participants in Study 4.2 selected ideas from a pre-generated set of ideas; all of these ideas were taken from previous studies, where I used the "education" topic. Half of the participants were instructed to select the most *creative* ideas, without further specification (creativity instructions), while the other participants were instructed to select "the best" ideas. Furthermore, half of the participants were instructed to give explicit quality ratings for all ideas before making their selection, while the other participants immediately made their selection (and rated the ideas afterwards).

There was high agreement between my participants and my raters with regard to the relative originality and feasibility of the ideas contained in the idea set. Further, participants with creativity instructions selected more original ideas than participants without creativity instructions, and these selected ideas also had a higher average originality than the ideas contained in the idea set, while the ideas selected by participants without creativity instructions were *less* original than the ideas in the idea set. Participants with creativity instructions selected less feasible ideas than participants without creativity instructions, but only when they had first rated the ideas. Among the participants who had first rated the ideas, the ideas selected by participants with creativity instructions were also less feasible than the idea set, whereas the ideas selected by participants without creativity instructions were more feasible than the idea set. Among participants who had not rated the ideas before making their selection, creativity instructions had no effect on the feasibility of the selected ideas.

Results from the postexperimental questionnaire showed a strong correlation between participants' tendency to select for originality and creativity. In contrast, participants' tendency to select feasible ideas was highly correlated with their tendency to select desirable or useful ideas. Creativity instructions increased participants' motivation to select creative and original ideas, while instructions to select "the best idea" increased participants' motivation to select feasible and useful ideas. Furthermore, the degree to which participants relied on originality/creativity in the one hand, and feasibility/usefulness on the other hand, were strongly negatively correlated.

Apparently, people are able to make a more creative selection when they are instructed to do so, but only when they do not perceive the selection criteria to be contradictory. Interestingly, the higher originality of ideas selected under creativity instructions was not necessarily accompanied by a lower feasibility. Thus, although people seem to perceive originality and feasibility to be incompatible criteria and tend to make a trade-off between these criteria, this trade-off need not be reflected in actual quality of the selected ideas.

Because participants in Study 4.2 only selected ideas from a pre-generated idea set, I conducted a new study to test whether the same effect would occur when participants made a selection from ideas they had generated in a brainstorming task. In Study 4.3, therefore, participants both generated and selected ideas. Again, the brainstorming topic concerned possible improvements in education at the Department of Psychology. As manipulation of criteria, half of the participants were instructed to keep their experience as students in mind (relevance instructions) during idea generation and idea selection, while the other participants were instructed to keep originality in mind (originality instructions). Furthermore, I wanted to test whether explicitly manipulating the average originality of the available ideas would also affect the originality of the selected ideas. My previous studies had shown that the quality of selected ideas was not so much related to the absolute availability of creative ideas, but rather to the *average* quality of the available ideas. In this study, therefore, participants were provided with either a broad (improvements in education at the Department of Psychology) or a narrow (improvements in lectures at the Department of Psychology) brainstorming topic. In line with my findings from the Chapter 3, participants with a narrow problem generated ideas that were more original than those generated by participants with a broad problem, while productivity was not affected. The ideas generated by participants with originality instructions were marginally more original than those generated by participants with relevance instructions. However, there were no effects on the feasibility of the generated ideas. Further, the ideas selected by participants with originality instructions were more original than those selected by participants with relevance instructions, and more original than their generated ideas; this latter effect did not occur for participants with relevance instructions. Interestingly, there were no effects on the feasibility of the selected ideas. Finally, problem scope did not affect selection performance.

These results show that a focus on originality can increase the originality of generated and selected ideas, and can bring selection effectiveness above the level of chance for originality, without decreasing feasibility. However, this higher performance was accompanied by lower satisfaction with the quality of the generated ideas, which indicates that an originality focus did not change participants' actual preferences. Although the quality of the selected ideas in my previous studies was consistently related to the average quality of the generated ideas, explicitly manipulating the average quality of the generated ideas in this study did not affect the quality of the selected ideas.

The main conclusion to be drawn from Chapter 4 is that selection criteria can improve selection performance, but only if these are not perceived to be contradictory. People make a trade-off between originality and feasibility (favoring feasibility), and find it hard to combine these two dimensions. Because people have a strong tendency to focus on selecting feasible and desirable ideas, increasing the average originality of the available ideas and inducing deeper processing of the available ideas does not lead to the selection of better ideas.

## 5.3 Implications

The rationale behind the brainstorming procedure is that it is useful to generate many ideas, because this increases the availability of creative ideas, and therefore also increases the chance of the best possible idea being found and implemented. The main question in this dissertation was whether this assumption would be supported by empirical data. Specifically, the questions I studied were: (1) to what degree does productivity contribute to the selection of creative ideas, (2) is the relationship between quantity and quality in idea generation best described by a Darwinian approach, and (3) what factors contribute to (or hinder) the selection of creative ideas?

### *5.3.1 The Benefits of Being Productive*

The results reported in this dissertation clearly demonstrate that a higher productivity, although it leads to a higher availability of high-quality ideas, does *not* make it more likely that a creative idea is selected. Firstly, the higher productivity of nominal groups as compared to interactive groups did not lead to their selecting better ideas, despite the fact that this higher productivity was also associated with a higher availability of high-quality ideas. Secondly, productivity was unrelated to the quality of the selected ideas; thus, regardless of whether participants brainstormed individually or in a group, it was not the case that the most productive participants selected the best ideas. Thirdly, despite the correlation between the average quality of the generated ideas and the average quality of the selected ideas, manipulations that affected the average originality of the generated ideas had no effect on the originality of the selected ideas.

*Is Brainstorming Useless?*

These results would seem to argue strongly against the reasoning behind brainstorming. According to this reasoning, increasing productivity is supposed to increase the chance that high-quality ideas are selected. Instead, my findings indicate that productivity appears to be simply *irrelevant* for the quality of selected ideas. The question thus arises whether brainstorming contributes to the quality of selected ideas or solutions at all. However, it should be emphasized that none of my studies made a comparison between brainstorming and non-brainstorming conditions. Thus, although there appears to be no relation between productivity and the quality of selected ideas under brainstorming conditions, this does not mean that brainstorming itself has no effect on the quality of selected ideas. Nevertheless, my results do show that such a (hypothetical) effect is very unlikely to be mediated by productivity, which again casts doubt upon the usefulness of brainstorming. If the generation of a large number of creative ideas does not improve the quality of the ideas that get selected, it seems that people might do just as well simply trying to generate one or two very good ideas. This is in line with a conclusion drawn by Johnson (1972), who wrote that "it appears that the standard condition of asking the subject to write one solution to a problem is a hard one to beat" (p. 316).

*The Recommendation Against Group Brainstorming.*

The issue could be approached from another direction. For years, brainstorming researchers have argued against the use of interactive groups, because of the well-established productivity loss. If quantity is the final goal of brainstorming, nominal groups indeed seem to be better suited to attain that goal (with the possible exception of some forms of electronic brainstorming; see Dennis and Williams, 2003 for an overview). However, if quantity is only a step on the way to the selection and implementation of creative solutions, my results suggest that interactive groups are not at a particular disadvantage. The ideas selected by interactive groups in my studies were no worse than those selected by nominal groups. Thus, although my findings raise serious doubts about the contribution that brainstorming can make to the selection and implementation of creative solutions, the traditional recommendation against group brainstorming does not seem to be strongly supported, either. With regard to the quality of selected ideas, it does not really seem to matter if one uses interactive or nominal groups. Further, the use of interactive groups may fulfill other functions (for example, people find group brainstorming more enjoyable; also see Sutton and Hargadon, 1996). Hence, if one lets go of the strict focus on productivity, there may even be reasons to prefer interactive brainstorming groups over nominal groups.

### *5.3.2 Quantity and Quality in Idea Generation*

In line with previous findings, my results support the notion that a higher productivity is linearly associated with a higher availability of high-quality ideas. Moreover, productivity appears to be unrelated to the average quality of the generated ideas. Both findings are in accordance with a strict Darwinian view of idea generation: If idea generation is a more or less random process, the distribution of high-quality ideas across one's total output is random as well.

In contrast, I argued that the chance of generating a creative idea is not fixed, but increases with deeper exploration of relevant domain knowledge. In accordance with this reasoning, I found that a higher relative productivity within a particular subcategory of the brainstorming topic was associated with a higher average originality of the ideas within that subcategory, and that a narrow brainstorming topic caused participants to generate ideas of higher originality than a broad brainstorming topic. These findings are incompatible with the view of idea generation as an entirely stochastic process. If the chance of generating an original idea is fixed, the average originality of generated ideas should be independent of productivity, just as the proportions of heads and tails should be independent of the number of times one throws a coin. Instead, if the chance of generating an original ideas depends on the way in which the relevant domain knowledge is accessed, the average originality of generated ideas could increase through deeper exploration of domain knowledge. My results show that this is indeed the case.

*Cognitive Constraints*

My finding that generating ideas within a more narrow domain increased the originality of the generated ideas is in line with Osborn's (1963) recommendation to use specific, rather than vague problems for brainstorming. In fact, Osborn specifically recommends the decomposition of a brainstorming problem into more specific sub-problems. Previous research (e.g., Dennis et al., 1996; Coskun et al, 2000) has shown that this increases productivity; my results suggest that it could also increase the average quality of the generated ideas. Thus, constraining the way in which people approach a brainstorming problem does seem to be beneficial for creative performance.

As I also noted in Chapter 3 (also see Finke, Ward, & Smith, 1992, p. 32), there seems to be something paradoxical about a beneficial effect of restrictions on creativity. Creative behavior is traditionally associated with a 'free' style of thinking, not with constraints and limitations. Indeed, in the context of motivation, this view seems to be correct: People are more creative when they do not feel controlled (see Collins & Amabile, 1999, for an overview). However, since creative behavior by definition means behaving in an unusual manner, there is at least one unavoidable constraint associated with creativity: the requirement to let go of routine behavior and habitual modes of thinking. Cognitive constraints, such as those described in Chapter 3 and in Study 4.3, are tools that can help to fulfill that requirement.

### *5.3.3 The Bias Against Originality*

My results consistently show that, even after a brainstorming session, in which the generation of creative ideas is encouraged, people do not select very original and unusual ideas, unless they are explicitly instructed to do so. Instead, people focus on the perceived feasibility of their ideas and try to select ideas that they feel are important and useful. Although this does not necessarily rule out the selection of original ideas, people tend to act as if it does. More specifically: People tend to treat originality and feasibility as selection criteria that are strongly negatively correlated, while this may not always be the case for the actual quality of the available ideas. In general, people seem to display a bias against originality: They prefer feasible ideas instead, and are less satisfied with their selection when they are instructed to select original ideas.

It is clear, then, that the creative attitude and atmosphere which the brainstorming instructions are intended to establish, do not carry over to subsequent idea selection. This switch is particularly visible in the questionnaire items from Study 4.3: With regard to idea generation, there was a positive correlation between participants' tendency to be original and creative and their satisfaction with their performance. When it came to idea selection, however, participants who tried to select original and creative ideas were *less* satisfied with their performance. Apparently, people do not set the same standards for their own performance in idea generation and selection, regardless of what they are instructed to do. Nevertheless, my results show that people *are* able to recognize and select creative ideas, even if they do not feel that these ideas are really the best ones.

*Realistic Problems and the Preference for Feasibility*

Actually, what I call a bias against originality might be just as well described as a bias in favor of feasibility, accompanied by the belief that feasibility and originality are incompatible. In general, my participants tended to favor practical and uncreative solutions for every-day problems. Participants who were induced to give priority to more creative ideas consequently were less satisfied with their selection. This may partly have been due to the brainstorming topics I used in my experiments (education and health): These were highly realistic and personally relevant to my participants, and research has shown that people tend to be less creative with relevant and realistic problems (e.g., Dillon, Graham & Aidells, 1972; Harari & Graham, 1975).

The choice for realistic problems was a deliberate one. While it would be interesting to study idea selection with a less realistic or relevant problem, such as the "thumbs" problem ("What would be the advantages and disadvantages if people were to grow an extra thumb?"), this would not be very informative with regard to real-world brainstorming. The question therefore arises whether one can make people actually *prefer* creative solutions to a relevant problem. Simply stressing the importance of "being creative" does not seem to be sufficient. Furthermore, as Study 4.3 demonstrates, narrowing down the brainstorming topic seems to help in the idea generation stage, but not in the selection stage. An alternative possibility would be to frame the problem differently. For example, my results in Chapter 4 show that my participants had a strong preference for generating and selecting solutions for existing problems, rather than innovative ideas. This tendency to adapt, rather than innovate (e.g., Kirton, 1994), may have been due to my explicitly asking for ideas about "improvements" in education. An interesting extension of the studies reported here would be to instruct participants to generate ideas about possible *innovations* (as opposed to improvements) in education at the Department of Psychology. Phrasing the brainstorming topic in such a way might motivate participants to approach the problem more creatively and to focus less on generating and selecting practical solutions.

### *5.3.4 Practical Recommendations*

The results of these studies give rise to some practical suggestions. These should of course not be seen as ready-made solutions, but if applied judiciously they can contribute to more effective creative problem-solving.

*Problem Finding*

Firstly, with regard to idea generation, I should stress the importance of what Osborn (1963) called *problem finding*, the stage in which the actual problem or brainstorming topic is defined and phrased. My results show that properties of the brainstorming problem, and the way in which this problem is approached, can significantly affect the quality of the generated ideas. Specifically, one should avoid problem definitions that allow people to rely on habitual modes of thought and to give routine responses. Even worse are problem definitions that allow for simple retrieval of existing solutions or ideas. Instead, one should aim for a problem definition that encourages deep exploration and creative application of the available knowledge. Using 'narrow' problem definitions (possibly going through various subcategories of an overall problem sequentially, as recommended by other researchers) may help to avoid the pitfalls of habitual thinking.

*Idea Selection As a Distinct Process*

Secondly, it is important to pay attention to idea selection as a distinct process with its own problems and challenges. My results show that the traditional recommendation to separate idea generation and selection is a good one, especially in interactive groups. My results further suggest that idea generation and idea selection are not affected by the same factors, and this in turn suggests that interventions to improve the effectiveness of idea generation are, by themselves, not likely to improve the effectiveness of idea selection. Thus, it is probably wise to adopt an explicit multi-stage view of the creative process. One may wonder how to enhance creativity, but this will do little good if no attention is paid to the question which specific stage of the creative process, or which specific kind of creative behavior, should be enhanced.

*Selection Criteria*

Thirdly, with regard to selection criteria, it may be worthwhile to spend some time on clarification and explanation of the different possible selection criteria and the way in which these relate to each other, before idea selection is actually undertaken. I have found that people's perceptions of different quality dimensions are not always in accordance with the actual state of affairs. People generally seem to view originality and feasibility as strongly negatively correlated, although this need not be the case. My results suggest that these perceptions can become a self-fulfilling prophecy, preventing an optimal selection. Hence, attention should be paid to the fact that these criteria can very well be combined into an effective selection. In this context, it is also important to recall the fourth brainstorming rule: "combine and improve." Ideas that are generated in a brainstorming session should not be treated as fixed and immutable entities which people must either accept or reject. An original, but infeasible idea may become more feasible with adaptation, or it may be the starting point for the development of new, more feasible ideas. It is important not to reject ideas out of hand because they are too unusual or seem too unfeasible. Thus, although it may seem strange, some form of deferment of judgment may be important in idea selection as well as in idea generation.

## 5.4 Limitations and Future Research

All answers found through scientific study raise new questions, and the answers provided in this dissertation are no exception. In this paragraph, some of these questions will be addressed.

### *5.4.1 Group Effectiveness in Idea Selection*

One of the main themes running through this dissertation is the question of group effectiveness in idea selection. Although my results are clear (interactive groups are neither worse, nor better than nominal groups when it comes to idea selection), many questions remain. For example, given my findings that selection criteria can improve the effectiveness, it would be interesting to study the effects of selection criteria in interactive groups. As mentioned in Chapter 4, research by Laughlin and colleagues (e.g., Laughlin, Bonner, & Altermatt, 1998; Laughlin & Shippy, 1983; Laughlin, VanderStoep, & Hollingshead, 1991) has shown that groups can outperform individuals on intellective tasks. Although idea selection is a judgmental task (i.e., a task with no demonstrably correct solution), the presence of explicit selection criteria that all group members agree about may give the selection task a more intellective quality, which may give groups an advantage. However, it is hard to predict what would actually happen. For example, if the group members have a clear preference for feasible and useful ideas, group polarization (e.g., Myers & Lamm, 1976) could exacerbate the existing bias against originality, obstructing the selection of creative ideas. Alternatively, the group members need not be unanimous in their preferences; one or a few group members may prefer more creative and unusual approaches to a problem. Research on minority influence suggests that this may have a significant impact on the group's creative behavior (e.g., De Dreu & West, 2001; Nemeth, 1995). Thus, it is not clear what to expect, and further research on creative performance in interactive groups should certainly address the issue of group processes and group effectiveness in idea selection.

### *5.4.2 Goals, Motives and Intentions*

At several points in this dissertation, I have touched upon the issue of what goals people have in idea generation and selection. My results suggest that the bias against the selection of original ideas is at least partly due to a tendency to select solutions to existing problems, rather than innovations. This in turn suggests an important role for motivation, goals and intentions. Most of the work done in this area has focused on the issue of intrinsic motivation (the motivation to engage in an activity for its own sake) and creativity (e.g., Collins & Amabile, 1999; Eisenberger & Shanock, 2003), and on the effect of various types of rewards and instructions on creative behavior (e.g., Eisenberger & Cameron, 1996; Eisenberger & Rhoades, 2001; Shalley, 1991). In contrast, the question of which particular goals people want to achieve in a creative task has not been studied extensively. This question is particularly important in the context of organizational creativity and innovation. One interesting avenue of future research would be to forge a link between creativity research and such concepts as regulatory focus (Higgins, 1997; see Friedman & Förster, 2001 for an example of such a link) and deliberative and implemental mind-sets (Gollwitzer, Heckhausen, & Steller, 1990).

### *5.4.3 Problem and Task Properties in Creativity Research*

I have already mentioned (Chapter 4, and earlier in this chapter) the importance of studying the effects that problem definitions and interpretations may have in creative performance. Stein (1975), in a critical review of the brainstorming literature, notes that "[t]here seems to be little attention paid to the basic structural characteristics of the problems used in the studies of brainstorming, or for that matter to the characteristics of problems in any study of the procedures of creative problem solving" (p. 68), and this has not really changed in the past thirty years.

The question of how problem properties affect creativity is conceptually very close to the topic of problem finding. After all, the properties of a problem are at least partly shaped by the way in which it is phrased and defined. There has been research addressing the problem-finding stage in creativity (e.g., Getzels, 1975; Getzels & Csikszentmihalyi, 1976; Runco, 1994), but none of these studies have addressed brainstorming as such. This is a gap in the literature that should be filled. One cannot safely generalize the results of brainstorming studies using only a limited number of problems and topics, if one does not know to what extent the results may have been affected by the specific topics used. Furthermore, the issue of problem properties is interesting in its own right, both from a theoretical and an applied perspective. As mentioned in Chapter 1, Osborn (1963) explicitly included problem-finding in his three-stage model of the creative process. Thus, if brainstorming research is to take more account of the context in which brainstorming takes place (about which more in the next paragraph), problem finding is one of the issues that should be addressed.

### *5.4.4 The Context of Brainstorming*

It is an often-heard complaint that experimental research on brainstorming and idea generation pays too little attention to the (organizational) context of creative behavior. Especially in organizational psychology, where creativity is usually studied in the context of innovation, researchers advocate a broader and more inclusive approach to creativity. For example, West (2002) encourages the development of "models … which reciprocally link the organisation [sic] and wider environment of the group to characteristics of the group and its task, and to group creativity and innovation implementation" (p. 378). The emphasis of such models (e.g., Taggar, 2002; Woodman, Sawyer, & Griffin, 1993) usually is on the external environment in which creative behavior takes place. Knowledge about the moderating effects of contextual factors is important, especially if creativity research is to yield recommendations for organizational practice. However, attention should also be devoted to what could be called the *internal context* of creativity: the different stages of creative behavior and the way in which these interact. As I noted in Chapter 1, idea generation does not happen in a vacuum; ideas are generated for a reason, and in response to a problem or question. Different factors in the external environment are likely to affect different stages in the creative process in different ways. For example, factors at the organizational level (such as organizational goals) will probably have a stronger influence on the recognition and formulation of problems and on the implementation of ideas and solutions than on the actual generation of new ideas. This latter stage of the creative process is more likely to be affected by individual and team level factors (e.g., Taggar, 2002).

The studies reported in this dissertation are a starting point for a more inclusive approach to applied creativity, but of course only to a certain degree. Most notably, I did not pay attention to specific implementation intentions. Research has shown that people generate fewer ideas when they expect these ideas to be actually used (Harari & Graham, 1975), and it is very likely that such expectations would also affect idea selection.

The issue of context is also important in light of the fact that my studies were all laboratory experiments, with mostly undergraduate Psychology students as participants. If context is so important to creative behavior, should this not be more strongly reflected in my methodology? I would answer this question in the negative. One reason for this is that I think that my use of Psychology students as participants is actually quite realistic with regard to at least one of the brainstorming topics that I used: the education topic. Presumably, brainstorming sessions in organizations are usually conducted with participants who have at least some personal connection to the topic at hand, who have an interest in the problem being solved, and who have extensive knowledge of the issues involved. This is exactly the situation that existed with my student participants.

More importantly, research has shown that laboratory experiments in psychology have a higher external validity than is often assumed (e.g., Anderson, Lindsay & Bushman, 1999). Although there are many important differences between group (and individual) behavior inside and outside of the psychological laboratory, the suggestion that laboratory experiments are unrealistic and meaningless is gratuitous if it is not accompanied by an explanation of why these experiments fail to address the specific questions under study.

Nevertheless, there are many differences between 'real' groups and teams in organizations and my experimental groups (such as the 'ad hoc'-character of my experimental groups, as opposed to the shared history and future of teams in organizations), and these differences may have affected my results. Far from disqualifying my results, I think that these differences suggest interesting avenues for further research.

## 5.5 Conclusion

The results of my studies are a mixed bag for brainstorming researchers and practitioners. On the one hand, brainstorming really does not seem to be very useful for the creative process. On the other hand, the productivity loss of interactive brainstorming groups does not seem to be as big a problem as is traditionally assumed. Furthermore, there are ways to improve the quality of ideas generated in a brainstorming session, but this does not affect the outcome of idea selection. Idea selection can be improved by using the right selection criteria, but at the cost of participant satisfaction. The validity of Osborn's (1963) theory and recommendations remains somewhat doubtful, but it is clear that there is more to brainstorming than productivity and productivity losses.

This dissertation started out with a remark about the supposedly elusive and 'unscientific' nature of creativity. Sternberg and Lubart (1999), in their introductory chapter of "The handbook of creativity," state, almost indignantly, that this prejudice has impeded the progress of creativity research:

> *"The mystical approaches to the study of creativity have probably made it harder for scientific psychologists to be heard. Many people seem to believe (…) that creativity is something that just doesn't lend itself to scientific study, because it is a spiritual process. We believe that it has been hard for the scientific approach to shake the deep-seated view of some people that, somehow, psychologists are treading where they should not" (p. 5).*

They then go on to write:

> *"Equally damaging to the scientific study of creativity, in our view, has been the takeover of the field, in the popular mind, by those who follow what might be referred to as a pragmatic approach. Those taking this approach have been concerned primarily with developing creativity, secondarily with understanding it, but almost not at all with testing the validity of their ideas about it. (…) Of course, techniques can work in the absence of psychological theory or validation. But the effect of such approaches is often to leave people associating a phenomenon with commercialization and to see it as less than a serious endeavor for psychological study" (pp. 5-6).*

Sternberg and Lubart explicitly mention Osborn (along with Edward De Bono) as one of the best-known proponents of this "pragmatic approach." Indeed, Osborn was mainly concerned with practical recommendations for commercial organizations, not with scientific research. However, one should keep in mind that Osborn published his book in a time when only a small minority of psychologists was prepared to undertake the "serious endeavor" of creativity research. Moreover, the fact that Osborn was totally wrong in his prediction regarding group productivity has led to an important series of studies, without which we would know much less about idea generation and creativity. Thus, the word "damaging" seems quite unfair.

Rather than ignoring or running down the theories and recommendations developed in the pragmatic tradition, psychologists should feel inspired to test them and, if possible, improve upon them. If practitioners like Osborn and De Bono fail to adequately test their theories, that should be all the more reason for psychologists to do so instead. Sometimes research will wholly or partly confirm these theories; sometimes it will reveal a completely different picture. In all cases, new knowledge will be the result, and psychology only stands to gain. Of course, a critical view remains essential, especially in light of the present-day profusion of supposedly creativity-enhancing techniques and procedures. If we are to believe the many websites, brochures and books on creativity, being creative is a simple thing if one only follows the right rules. Although the studies reported in this dissertation may have done something to dispel that notion, there apparently exists a persistent need for clear-cut and uncomplicated advice. If there is one simple recommendation that can be distilled from my results, it is this: Creativity practitioners *and* researchers would do well to worry less about quantity, and more about quality – before, during, and after idea generation.

# References


AMABILE, T. M. (1979). Effects of external evaluation on artistic creativity. *Journal of Personality and Social Psychology, 37*, 221-233.

AMABILE, T. M. (1996). *Creativity in context: Update to "The social psychology of creativity."* Boulder, CO, US: Westview Press.

ANDERSON, C. A., LINDSAY, J. J., & BUSHMAN, B. J. (1999). Research in the psychological laboratory: Truth or triviality? *Current Directions in Psychological Science, 8*, 3-9.

ASCH, S. E. (1955). Opinions and social pressure. *Scientific American, 193(5)*, 31-55.

BAER, J. (2003). Evaluative thinking, creativity, and task specificity: Separating wheat from chaff is not the same as finding needles in haystacks. In M. A. Runco (Ed.)*, Critical creative processes* (pp. 129-151). Cresskill, NJ, US: Hampton Press, Inc.

BARGH, J. A., BOND, R. N., LOMBARDI, W. J., & TOTA, M. E. (1986). The additive nature of chronic and temporary sources of construct accessibility. *Journal of Personality and Social Psychology, 50*, 869-878.

BARON, R. M., & KENNY, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173-1182.

BARTIS, S., SZYMANSKI, K., & HARKINS, S. (1988). Evaluation and performance: A two-edged knife. *Personality and Social Psychology Bulletin, 14*, 42-251.

BASADUR, M. (1995). Optimal ideation-evaluation ratios. *Creativity Research Journal, 8,* 63-75.

BEACH, L. R. (1993). Broadening the definition of decision making: The role of prechoice screening of options. *Psychological Science, 4,* 215-220.

CAMACHO, L. M., & PAULUS, P. B. (1995). The role of social anxiousness in group brainstorming. *Journal of Personality and Social Psychology, 68*, 1071-1080.

CAMPBELL, D. T. (1960). Blind variation and selective retentions in creative thought as in other knowledge processes. *Psychological Review, 67*, 380-400.

CHARTRAND, T. L. (2005). The role of conscious awareness in consumer behavior. *Journal of Consumer Psychology, 15*, 203-210.

CHRISTENSEN, P. R., GUILFORD, J. P., & WILSON, R. C. (1957). Relations of creative responses to working time and instructions. *Journal of Experimental Psychology, 53*, 82-88.

CICCHETTI, D. V., & SPARROW, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency, 86*, 127-137.

COLLAROS, P., & ANDERSON, L. R. (1969). Effects of perceived expertness upon creativity of members of brainstorming groups. *Journal of Applied Psychology, 53*, 159-163.

COLLINS, M. A., & AMABILE, T. M. (1999). Motivation and creativity. In R. J. Sternberg (Ed.), *Handbook of Creativity* (pp. 297-312). Cambridge, UK: Cambridge University Press.

COSKUN, H., PAULUS, P. B., BROWN, V., & SHERWOOD, J. J. (2000). Cognitive stimulation and problem presentation in idea-generating groups. *Group Dynamics: Theory, Research and Practice, 4*, 307-329.

DE DREU, C. K. W., & WEST, M. A. (2001). Minority dissent and team innovation: The importance of participation in decision-making. *Journal of Applied Psychology, 68*, 1191-1201.

DENNIS, A. R., VALACICH, J. S., CONNOLLY, T., & WYNNE, B. (1996). Process structuring in group brainstorming. *Information Systems Research, 7,* 268-277.

DENNIS, A. R., & WILLIAMS, M L. (2003). Electronic brainstorming: Theory, research and future directions. In P. B. Paulus, & B. A. Nijstad (Eds.), *Group creativity: Innovation through collaboration* (pp. 160-178). Oxford, UK: Oxford University Press.

DENTLER, R. A., & MACKLER, B. (1964). Originality: Some social and personal determinants. *Behavioral Science, 9*, 1-7.

DIEHL, M.(1991). *Kollektive Kreativität: Zur Quantität und Qualität der Ideenproduktion in Kleingruppen* (Collective creativity: On quantity and quality of idea production in small groups). Habilitationschrift, University of Tübingen, Germany.

DIEHL, M., & STROEBE, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology, 53*, 497-509.

DIEHL, M., & STROEBE, W. (1991). Productivity loss in idea-generating groups: Tracking down the blocking effect. *Journal of Personality and Social Psychology, 61*, 392-403.

DILLON, P. C., GRAHAM, W. K., & AIDELLS, A. L. (1972). Brainstorming on a "hot" problem: Effects of training and practice on individual and group performance. *Journal of Applied Psychology, 56*, 487-490.

DUGOSH, K. L., PAULUS, P. B., ROLAND, E. J., & YANG, H. C. (2000). Cognitive stimulation in brainstorming. *Journal of Personality and Social Psychology, 79*, 722-735.

EISENBERGER, R., & CAMERON, J. (1996). Detrimental effects of reward: Reality or myth? *American Psychologist, 51*, 1153-1166.

EISENBERGER, R., & RHOADES, L. (2001). Incremental effects of reward on creativity. *Journal of Personality and Social Psychology, 81*, 728-741.

EISENBERGER, R., & SHANOCK, L. (2003). Rewards, intrinsic motivation, and creativity: A case study of conceptual and methodological isolation. *Creativity Research Journal, 15*, 121-130.

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

FEIST, G. J., & RUNCO, M. A. (1993). Trends in the creativity literature: An analysis of research in the *Journal of Creative Behavior* (1967-1989). *Creativity Research Journal, 6*, 271-286.

FINKE, R. A., WARD, T. B., & SMITH, S. M. (1992). *Creative cognition: Theory, research, and applications*. Cambridge, MA, US: The MIT Press.

FRIEDMAN, R. S., FISHBACH, A., FÖRSTER, J., & WERTH, L. (2003). Attentional priming effects on creativity. *Creativity Research Journal, 15*, 277-286.

FRIEDMAN, R. S., & FÖRSTER, J. (2001). The effects of promotion and prevention cues on creativity. *Journal of Personality and Social Psychology, 81*, 1001-1013.

FURR, R. (2004). Interpreting effect sizes in contrast analysis. *Understanding Statistics, 3*, 1-25.

GALLUPE, R. B., BASTIANUTTI, L. M., & COOPER, W. H. (1991). Unblocking brainstorms. *Journal of Applied Psychology, 76*, 137-142.

GETZELS, J. W. (1975). Problem-finding and the inventiveness of solutions. *Journal of Creative Behavior, 9*, 12-18.

GETZELS, J. W., & CSIKSZENTMIHALYI, M. (1976). *The creative vision: A longitudinal study of problem finding in art.* New York, NY, US: Wiley.

GUILFORD, J. P. (1950). Creativity. *American Psychologist, 5*, 444-454.

GOLLWITZER, P. M., HECKHAUSEN, H., & STELLER, B. (1990). Deliberative and implemental mind-sets: Cognitive tuning toward congruous thoughts and information. *Journal of Personality and Social Psychology, 59*, 1119-1127.

HARARI, O., & GRAHAM, W. K. (1975). Tasks and task consequences as factors in individual and group brainstorming. *Journal of Social Psychology, 95,* 61-65.

HARRINGTON, D. M. (1975). Effects of explicit instructions to "be creative" on the psychological meaning of divergent thinking test scores. *Journal of Personality, 43,* 434-454.

HASTIE, R., & DAWES, R. M. (2001). *Rational choice in an uncertain world: The psychology of judgment and decision making*: Thousand Oaks, CA, US: Sage Publications Inc.

HELLER, D., LEVIN, I. P., & GORANSSON, M. (2002). Selection of strategies for narrowing choice options: Antecedents and consequences. *Organizational Behavior and Human Decision Processes, 89,* 1194-1213.

HENNESSEY, B. A. (1989). The effect of extrinsic constraints on children's creativity while using a computer. *Creativity Research Journal, 2*, 151-168.

HIGGINS, E. T. (1989). Knowledge accessibility and activation: Subjectivity and suffering from unconscious sources. In J. A. Bargh, & J. S. Uleman (Eds.), *Unintended thought* (pp. 75-123). New York, NY, US: Guilford Press.

HIGGINS, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In A. W. Kruglanski, & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (pp. 133-168). New York, NY, US: Guilford Press.

HIGGINS, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*, 1280-1300.

JANSSON, D. G., & SMITH, S. M. (1991). Design Fixation. *Design Studies, 12*, 3-11.

JOHNSON, D. M. (1972). *A systematic introduction to the psychology of thinking.* New York, NY, US: Harper & Row.

KAPLAN, M. F., & MILLER, C. E. (1987). Group decision making and normative versus informational influence: Effects of type of issue and assigned decision rule. *Journal of Personality and Social Psychology, 53,* 306-313.

KARAU, S. J., & KELLY, J. R. (1992). The effects of time scarcity and time abundance on group performance quality and interaction process. *Journal of Experimental Social Psychology, 28,* 542-571.

KIRTON, M. J. (Ed.) (1994). *Adaptors and innovators: Styles of creativity and problem solving.* London: Routledge.

LAMM, H., & TROMMSDORFF, G. (1973). Group versus individual performance on tasks requiring ideational proficiency (brainstorming): A review. *European Journal of Social Psychology, 3,* 361-388.

LAUGHLIN, P. R. (1980). Social combination processes of cooperative problem-solving groups on verbal intellective tasks. In M. Fishbein (Ed.), *Progress in Social Psychology* (pp. 127-155). Hillsdale, NJ, US: Lawrence Erlbaum.

LAUGHLIN, P. R., & SHIPPY, T. A. (1983). Collective induction. *Journal of Personality and Social Psychology, 45,* 94-100.

LAUGHLIN, P. R., BONNER, B. L., & ALTERMATT, T. W. (1998). Collective versus individual induction with single versus multiple hypotheses. *Journal of Personality and Social Psychology, 75,* 1481-1489.

LAUGHLIN, P. R., VANDERSTOEP, S. W., & HOLLINGSHEAD, A. B. (1991). Collective versus individual induction: Recognition of truth, rejection of error, and collective information processing. *Journal of Personality and Social Psychology, 61,* 50-67.

LEVIN, I. P., HUNEKE, M. E., & JASPER, J. D. (2000). Information processing at successive stages of decision making: Need for cognition and inclusion-exclusion effects. *Organizational Behavior and Human Decision Processes, 82,* 171-193.

LUCHINS, A. S. (1942). Mechanization in problem solving--the effect of Einstellung. *Psychological Monographs, 54(6)*: 95.

MACKINNON, D. P., LOCKWOOD, C. M., HOFFMAN, J. M., WEST, S. G., & SHEETS, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*, 83-104.

MACKINNON, D. P., WARSI, G., & DWYER, J. H. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research, 30,* 41-62.

MANSKE, M. E., & DAVIS, G. A. (1968). Effects of simple instructional biases upon performance in the unusual uses test. *Journal of General Psychology, 79,* 25-33.

MCGRAW, K. O., & WONG, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*, 30-46.

MEADOW, A., PARNES, S. J., & REESE, H. (1959). Influence of brainstorming instructions and problem sequence on a creative problem solving test. *Journal of Applied Psychology, 43*, 413-416.

MERRIAM-WEBSTER ONLINE DICTIONARY (n.d.). Retrieved August 21, 2005, from http://www.m-w.com/

MULLEN, B., JOHNSON, C.,& SALAS, E. (1991). Productivity loss in brainstorming groups: A meta-analytic integration. *Basic and Applied Social Psychology, 12*, 3-24.

MYERS, D G., & LAMM, H. (1976). The group polarization phenomenon. *Psychological Bulletin, 83*, 602-627.

NEMETH, C. J. (1995). Dissent as driving cognition, attitudes and judgments. *Social Cognition, 13*, 273-291.

NIJSTAD, B. A., & DE DREU, C. K. W. (2002). Creativity and group innovation. *Applied Psychology: An International Review, 51*, 400-406.

NIJSTAD, B. A., STROEBE, W., & LODEWIJKX, H. F. M. (2002). Cognitive stimulation and interference in groups: Exposure effects in an idea generation task. *Journal of Experimental Social Psychology, 38*, 535-544.

NIJSTAD, B. A., STROEBE, W., & LODEWIJKX, H. F. M. (2003). Production blocking and idea generation: Does blocking interfere with cognitive processes? *Journal of Experimental Social Psychology, 39*, 531-548.

OSBORN, A. F. (1953). *Applied imagination*. New York, NY, US: Charles Scribner's sons.

OSBORN, A. F. (1957). *Applied imagination, 2nd edition*. New York, NY, US: Charles Scribner's sons.

OSBORN, A. F. (1963). *Applied imagination, 3rd revised edition.* New York, NY, US: Charles Scribner's sons.

OXFORD ADVANCED LEARNER'S DICTIONARY (6th ed.) (2000). Oxford, UK: Oxford University Press.

PARKS, C. D., & COWLIN, R. (1995). Group discussion as affected by number of alternatives and by a time limit. *Organizational Behavior and Human Decision Processes, 62,* 267-275.

PARNES, S. J. (Ed.) (1992). *Source book for creative problem-solving: A fifty year digest of proven innovation processes.* Buffalo, NY, US: Creative Education Foundation Press.

PARNES, S. J., & MEADOW, A. (1959). Effects of "brainstorming" instructions on creative problem solving by trained and untrained subjects. *Journal of Educational Psychology, 50*, 171-176.

PAULUS, P. B., & DZINDOLET, M. T. (1993). Social influence processes in group brainstorming. *Journal of Personality and Social Psychology, 64*, 575-586.

PAULUS, P. B., DZINDOLET, M. T., POLETES, G., & CAMACHO, L. M. (1993). Perception of performance in group brainstorming: The illusion of group productivity. *Personality and Social Psychology Bulletin, 19*, 78-89.

PERKINS, D. N. (1981). *The mind's best work.* Cambridge, MA, US: Harvard University Press.

PUTMAN, V. L., & PAULUS, P. B. (2003). *Brainstorming, brainstorming rules and decision making.* Unpublished manuscript, University of Texas at Arlington.

RÉVÉSZ, G. (1952). *Talent en genie: Grondslagen van een psychologie der begaafdheid.* (Talent and genius: Foundations of a psychology of giftedness). Leiden: Brill.

RICKARDS, T. (1975). Brainstorming: An examination of idea production rate and level of speculation in real managerial situations. *R and D Management, 6,* 11-14.

RIETZSCHEL, E. F., NIJSTAD, B. A., & STROEBE, W. (in press). Productivity is not enough: A comparison of interactive and nominal brainstorming groups on idea generation and selection. *Journal of Experimental Social Psychology.*

RIETZSCHEL, E. F., NIJSTAD, B. A., & STROEBE, W. (2005a). Relative accessibility of domain knowledge and creativity: The effects of knowledge activation on the quantity and quality of generated ideas. *Manuscript submitted for publication.*

RIETZSCHEL, E. F., NIJSTAD, B. A., & STROEBE, W. (2005b). The selection of creative ideas after brainstorming: Overcoming the bias against originality. *Manuscript submitted for publication.*

RUNCO, M. A. (Ed.) (1994). *Problem finding, problem solving, and creativity.* Westport, CT, US: Ablex Publishing.

RUNCO, M. A., & CHARLES, R. E. (1993). Judgments of originality and appropriateness as predictors of creativity. *Personality and Individual Differences, 15,* 537-546.

SCHACTER, S. (1951). Deviation, rejection and communication. *Journal of Abnormal and Social Psychology, 46*, 190-207.

SHAFIR, E. (1993). Choosing versus rejecting: Why some options are both better and worse than others. *Memory and Cognition, 21,* 546-556.

SHALLEY, C. E. (1991). Effects of productivity goals, creativity goals, and personal discretion on individual creativity. *Journal of Applied Psychology, 76,* 179-185.

SHROUT, P. E., & FLEISS, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.

SIMON, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69*, 99-118.

SIMONTON, D. K. (1997). Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychological Review, 104*, 66-89.

SIMONTON, D. K. (1999a). Creativity as blind variation and selective retention: Is the creative process Darwinian? *Psychological Inquiry, 10*, 309-328.

SIMONTON, D. K. (1999b). *Origins of genius: Darwinian perspectives on creativity.* Oxford, UK: Oxford University Press.

SIMONTON, D. K. (2003). Scientific creativity as constrained stochastic behavior: The integration of product, person, and process perspectives. *Psychological Bulletin, 129*, 475-494.

SMITH, G. F. (1998). Idea-generation techniques: A formulary of active ingredients. *Journal of Creative Behavior, 32*, 107-133.

STASSER, G. (1999). The uncertain role of unshared information in collective choice. In L. L. Thompson, J. M. Levine, & D. M. Messick (Eds.), *Shared cognition in organizations* (pp. 49-69). Mahwah, NJ: Lawrence Erlbaum Associates.

STEIN, M. I. (1975). *Stimulating creativity: Group procedures.* Oxford, UK: Academic Press.

STERNBERG, R. J., & LUBART, T. I. (1999). The concept of creativity: Prospects and paradigms. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 3-15). Cambridge, UK: Cambridge University Press.

STROEBE, W., DIEHL, M., & ABAKOUMKIN, G. (1992). The illusion of group effectivity. *Personality and Social Psychology Bulletin, 18*, 643-650.

SUTTON, R. I., & HARGADON, A. (1996). Brainstorming groups in context: Effectiveness in a product design firm. *Administrative Science Quarterly, 41*, 685-718.

TAGGAR, S. (2002). Individual creativity and group ability to utilize individual creative resources: A multilevel model. *Academy of Management Journal, 45*, 315-330.

TAYLOR, D. W., BERRY, P. C., & BLOCK, C. H. (1958). Does group participation when using brainstorming facilitate or inhibit creative thinking? *Administrative Science Quarterly, 3*, 23-47.

TVERSKY, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79,* 281-299.

VINCENT, A. S., DECKER, B. P., & MUMFORD, M. D. (2002). Divergent thinking, intelligence, and expertise: A test of alternative models. *Creativity Research Journal, 14*, 163-178.

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

WARD, T. B., PATTERSON, M. J., SIFONIS, C. M., DODDS, R. A., & SAUNDERS, K. N. (2002). The role of graded category structure in imaginative thought. *Memory and Cognition, 30*, 199-216.

WEISBERG, R. W. (1995). Case studies of creative thinking: Reproduction versus restructuring in the real world. In S. M. Smith, & T. B. Ward (Eds.), *The creative cognition approach* (pp. 53-72). Cambridge, MA, US: The MIT Press.

WEISBERG, R. W. (1999). Creativity and knowledge: A challenge to theories. In R. J. Sternberg (Ed), *Handbook of creativity* (pp. 226-250). Cambridge, UK: Cambridge University Press.

WEST, M. A. (2002). Sparkling fountains or stagnant ponds: An integrative model of creativity and innovation implementation in work groups. *Applied Psychology: An International Review, 51,* 355-387.

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

WOODMAN, R. W., SAWYER, J. E., & GRIFFIN, R. W. (1993). Toward a theory of organizational creativity. *Academy of Management Review, 18*, 293-321.

YANIV, I., & SCHUL, Y. (1997). Elimination and inclusion procedures in judgment. *Journal of Behavioral Decision Making, 10,* 211-220.

# Appendix 1 – Coding Scheme for Education Topic

*How original (innovative, unusual) is the idea?*

*1: Not original*

Ideas that are mentioned very often, well-known complaints, ideas that concern measures that have already been taken, or ideas that involve only a minor change of the existing situation.
Examples: more interaction in classes; more information; no multiple choice tests

*2: Hardly original*

Ideas that are still mentioned fairly often, but concern somewhat more significant changes in the existing situation, such as more substantial changes in the educational program.
Examples: more practical assignments; more possibilities to participate in exchange programs; books in Dutch language

*3: Somewhat original*

These ideas can be called somewhat innovative, for example because they involve the use of new technology, or because they change the educational program through indirect means (and are therefore more unusual).
Examples: classes should be available on the Internet (through a webcam); the faculty building should be nicer; famous psychologists or psychiatric patients should be invited to give lectures

*4: Original*

These ideas are original and innovative, but not highly unusual or eccentric. These are often ideas that introduce a new element into the educational program. Most of these ideas are mentioned more than once, but not often.
Examples: give every student a notebook computer; create housing for all students next to the faculty; create a relaxation room with music and soft pillows

*5: Highly original*

These are rare, unusual or even radical ideas. These ideas may refer to innovative applications of existing technology, or even to non-existing technology.
Examples: follow psychologists with a camera and broadcast this on university TV channel; create a direct train connection to the faculty; hire aliens to teach about the psychology of *homo sapiens*

## *How feasible is the idea?*

*1: Not feasible*

These ideas are impossible to realize; for example, because the necessary means are not available, or because the implementation is a political matter and not something that can be arranged on a university level.

Examples: every student should receive a life-long scholarship; the Minister of education should be fired

*2: Hardly feasible*

These ideas are feasible in principle, but meet insurmountable practical problems.

Examples: students who get good grades, receive a refund of their tuition fee; all students are provided with a house by the university

*3: Somewhat feasible*

These are ideas that are feasible, but do meet practical difficulties; for example, because the costs are high, or because they involve major changes.

Examples: different tests should not be scheduled on the same day; there should be more study rooms in the faculty building

*4: Feasible*

These ideas are more easily implemented, but still involve fairly large changes or investments. This category also contains ideas that seem highly feasible, but are phrased very generally.

Examples: professors should receive more instruction on how to teach; there should be more emphasis on practical abilities

*5: Highly feasible*

These ideas can be implemented very easily, without large investments or complicated arrangements; for example, such an idea might propose a minor change of the existing situation. This category also contains ideas that do not constitute any change at all.

Examples: professors should make more use of PowerPoint presentations; there should be more information about the class and test schedules

# Appendix 2 – Category System for Health Topic

*Goals:*

01: maintain or improve physical condition (muscles, heart)

02: maintain or improve perception

03: optimize metabolism and avoid nutritional deficits

04: good posture, avoid skeletal problems

05: avoid or reduce over- and underweight

06: avoid external injuries

07: maintain or improve mental health

08: avoid physical strain

09: avoid toxic substances and radiation

10: prevention of illness

11: maintain dental health

*Means:*

1: nutrition

2: medicine and treatment

3: clothing

4: physical hygiene and care

5: avoiding, changing or seeking particular external influences

6: exercise

7: information and knowledge

8: lifestyle

9: social contacts

# Samenvatting (Dutch summary)

Een veelgebruikte methode om de productie van creatieve ideeën te bevorderen, is het zogeheten brainstormen, een procedure die in de jaren '50 door Alex Osborn is uiteengezet in zijn boek *Applied Imagination*. In een brainstormsessie is het de bedoeling dat deelnemers zoveel mogelijk originele ideeën bedenken, zonder zichzelf en elkaar te evalueren of te bekritiseren. De assumptie hierachter is dat *kwantiteit leidt tot kwaliteit*: het bedenken van *veel* ideeën wordt verondersteld automatisch samen te gaan met het bedenken van veel *goede* ideeën (een goed idee is een idee dat *origineel* en *uitvoerbaar* is). Hoe meer goede ideeën er bedacht worden, hoe groter de kans zou zijn dat er uiteindelijk een goed idee gekozen wordt om verder te ontwikkelen en te implementeren (zie **Hoofdstuk 1**). De centrale vraag van dit proefschrift is of deze assumptie juist is. Hoewel het bedenken van creatieve ideeën een noodzakelijke voorwaarde is voor het selecteren en implementeren van creatieve ideeën, is de combinatie van ideeënproductie en –selectie nog nauwelijks onderzocht. Bovendien is er in het bestaande brainstormonderzoek betrekkelijk weinig aandacht besteed aan de kwaliteit van de bedachte ideeën. Beide onderwerpen (de kwaliteit van ideeën, en de selectie van ideeën) staan centraal in dit proefschrift.

Hoewel brainstormen vaak in groepsverband gebeurt, is dat niet noodzakelijk. Bovendien heeft onderzoek aangetoond dat mensen die samen in een groep brainstormen (een *interactieve groep*) veel minder ideeën bedenken dan mensen die individueel brainstormen en wiens ideeën later bij elkaar opgeteld worden (een *nominale groep*). Groepsbrainstormen wordt daarom meestal door onderzoekers afgeraden. Een belangrijke vraag die hieruit voortvloeit, maar die nog nauwelijks is onderzocht, is of dit zogeheten *productiviteitsverlies* in interactieve groepen ook leidt tot de selectie van minder goede ideeën (zie **Hoofdstuk 2**). Als de kans op het kiezen van een goed idee inderdaad groter wordt naarmate er meer ideeën beschikbaar zijn om uit te kiezen, zouden we verwachten dat nominale groepen, die de meeste ideeën bedenken, ook de beste ideeën kiezen. Aan de andere kant is het mogelijk dat interactieve groepen een zorgvuldiger keuzeproces doormaken, en daardoor uiteindelijk net zulke goede of zelfs betere ideeën kiezen dan nominale groepen. Een vraag die hiermee samenhangt, is hoe *effectief* de selectie van creatieve ideeën eigenlijk is. Als mensen niet of nauwelijks in staat zijn hun beste ideeën te herkennen en te kiezen, heeft het bedenken van veel creatieve ideeën niet zoveel zin. Bovendien is het onduidelijk of een strikte scheiding tussen het bedenken en het kiezen van ideeën, zoals die in de traditionele brainstormprocedure gehanteerd wordt, wel zo belangrijk is.

In een experiment met interactieve en nominale brainstormgroepen (Studie 2.1) vond ik dat nominale groepen inderdaad meer ideeën, en meer goede ideeën, bedachten dan interactieve groepen, vooral wanneer de scheiding tussen het bedenken en kiezen van ideeën minder strikt werd aangehouden. Ondanks dit productiviteitsverschil was er geen verschil in de kwaliteit van de uiteindelijke keuze. Bovendien was het keuzeproces erg ineffectief: de kwaliteit van de gekozen ideeën was niet hoger of lager dan de gemiddelde kwaliteit van de bedachte ideeën; dat was zelfs het geval wanneer deelnemers expliciet de instructie kregen om ideeën te kiezen die origineel en uitvoerbaar waren (Studie 2.2). Met andere woorden: de keuze was niet beter dan een willekeurige greep geweest zou zijn. Bovendien was er geen relatie tussen de hoeveelheid bedachte ideeën en de kwaliteit van de gekozen ideeën. Kwantiteit lijkt dus maar in beperkte mate te leiden tot kwaliteit: wie meer ideeën bedenkt, bedenkt ook wel meer goede ideeën, maar kiest niet automatisch betere ideeën.

De relatie tussen kwantiteit en kwaliteit is nog niet uitgebreid onderzocht: brainstormonderzoekers richten zich meestal vooral op het aantal bedachte ideeën. De correlatie tussen het aantal bedachte ideeën en het aantal bedachte goede ideeën is over het algemeen zeer sterk, en onderzoekers nemen daarom doorgaans aan dat er weinig extra informatie te halen valt uit data over kwaliteit. Toch is het een interessante en belangrijke vraag hoe deze relatie nu in elkaar steekt. De meest simpele benadering is een *toevalsbenadering*: ieder idee dat bedacht wordt heeft een even grote kans om een goed idee te zijn, en daarom neemt het aantal goede ideeën proportioneel met het aantal bedachte ideeën toe. Deze benadering lijkt goed aan te sluiten bij de beschikbare gegevens over brainstormen. Hiertegenover staat de '*diepe verkenningsbenadering*,' die zegt dat niet ieder idee een even grote kans heeft om een goed idee te zijn, maar dat dit afhangt van de manier waarop mensen hun kennis gebruiken om ideeën te bedenken. Mensen zouden normaal gesproken geneigd zijn om hun kennis vooral oppervlakkig in te zetten, en zouden daardoor ideeën bedenken die niet erg origineel zijn. Naarmate ze hun kennis 'dieper' verkennen, en meer durven af te wijken van de gangbare benaderingen, wordt de kans op het bedenken van een echt origineel idee groter. Eén manier om mensen tot diepere verkenning van hun kennis te laten komen, is door ze te *beperken* in de manier waarop ze ideeën kunnen bedenken. Een beperking in de manier waarop een probleem benaderd kan worden, dwingt mensen om creatiever met hun kennis om te gaan.

De centrale vraag van **Hoofdstuk 3** was of het bedenken van ideeën in een brainstormsessie adequaat beschreven wordt door de toevalsbenadering, of dat er ook aanwijzingen gevonden kunnen worden voor positieve effecten van diepere verkenning. In de drie experimenten in dit hoofdstuk maakte ik gebruik van een *priming*-manipulatie. Alvorens ideeën te bedenken in een brainstormsessie, kregen deelnemers aan deze experimenten open vragen voorgelegd over een subonderwerp (voeding, hygiëne, of sport) van het uiteindelijke brainstormonderwerp (gezondheid). Het doel van deze open vragen was om de kennis die deelnemers over een subonderwerp hadden, te activeren. Ik verwachtte dat deelnemers die open vragen over een subonderwerp beantwoord hadden, verhoudingsgewijs meer en originelere ideeën over dat subonderwerp zouden bedenken dan deelnemers die geen open vragen hadden beantwoord. Dit bleek over het geheel genomen inderdaad het geval te zijn (Studie 3.1). Hoewel er geen verband was tussen de totale hoeveelheid ideeën die mensen bedachten en de gemiddelde originaliteit van die ideeën, was er wel een positief verband tussen het percentage ideeën dat mensen bedachten binnen een bepaalde subcategorie, en de gemiddelde originaliteit van de ideeën binnen die subcategorie. Dit effect van priming werd nog sterker wanneer mensen tijdens het brainstormen ideeën te zien kregen die over hetzelfde onderwerp gingen als de open vragen die ze beantwoord hadden (Studie 3.2). Ten slotte bleek deze priming-manipulatie ook te werken bij interactieve en nominale tweetallen (Studie 3.3). De effecten van priming op de originaliteit van de bedachte ideeën werkten echter niet door in de selectiefase: de priming-manipulatie had geen effect op de kwaliteit van de gekozen ideeën. Bovendien kozen interactieve en nominale groepen, net als in Studie 2.1, ideeën van gelijke kwaliteit. Deze ideeën hadden een hogere gemiddelde uitvoerbaarheid, maar een lagere gemiddelde originaliteit dan de bedachte ideeën. De selectie was dus nog steeds niet erg effectief.

De resultaten van de voorgaande studies laten zien dat de potentiële voordelen van het bedenken van creatieve ideeën tenietgedaan kunnen worden door een ineffectief selectieproces. Het is dus belangrijk om te weten welke factoren bijdragen aan de effectiviteit van ideeënselectie. In **Hoofdstuk 4** stel ik dat een effectieve ideeënselectie vereist dat (1) alle ideeën zorgvuldig in overweging worden genomen, en (2) de ideeën worden beoordeeld aan de hand van relevante kwaliteitscriteria. De ineffectieve ideeënselectie in eerdere studies zou te wijten kunnen zijn aan oppervlakkige of onvolledige beoordeling van de aanwezige ideeën, aan het gebruik van andere selectiecriteria dan originaliteit en uitvoerbaarheid, of aan een combinatie van deze twee factoren. In Studie 4.1 vond ik echter dat noch selectiecriteria (originaliteit en uitvoerbaarheid), noch een uitgebreide keuzeprocedure (het samenstellen van een keuzeset alvorens de definitieve keuze te maken), noch de combinatie van deze twee een effect hadden op de effectiviteit van de selectie: in geen van de experimentele condities week de gemiddelde kwaliteit van de gekozen ideeën af van de gemiddelde kwaliteit van de bedachte ideeën. In een volgende studie (Studie 4.2) bleek echter dat mensen die de instructie kregen om *creatieve* ideeën te kiezen, wèl ideeën kozen die een hogere gemiddelde originaliteit hadden dan de beschikbare ideeën. Uit een vragenlijst bleek ook dat deze deelnemers zich bij het maken van hun selectie sterker richtten op originaliteit, en minder sterk op uitvoerbaarheid, dan deelnemers die de instructie hadden gekregen om "de beste" ideeën te kiezen. Deze resultaten doen vermoeden dat de deelnemers in de studies 2.2 en 4.1 moeite hadden met de instructie om ideeën te kiezen die zowel origineel als uitvoerbaar waren. Met de juiste instructies blijkt het dus wel degelijk mogelijk te zijn om mensen een creatievere keuze te laten maken.

In een volgend experiment (Studie 4.3) bedachten deelnemers ideeën en maakten daarna een keuze uit hun eigen ideeën. De helft van de deelnemers kreeg de instructie om bij het bedenken en kiezen van ideeën zo origineel mogelijk te zijn. Deze deelnemers kozen inderdaad originelere ideeën dan deelnemers zonder originaliteitsinstructies, en deze gekozen ideeën hadden ook een hogere gemiddelde originaliteit dan de bedachte ideeën. Deze hogere originaliteit ging echter niet ten koste van de uitvoerbaarheid van de gekozen ideeën. Naast selectiecriteria werd in dit experiment ook de breedte van het probleem gemanipuleerd: de helft van de deelnemers bedacht ideeën over mogelijke verbetering in het onderwijs aan de afdeling psychologie (breed probleem), en de andere helft van de deelnemers bedacht ideeën over mogelijke verbeteringen in de *colleges* aan de afdeling psychologie (smal probleem). Beperking van de probleembreedte bleek wel een positief effect te hebben op de gemiddelde originaliteit van de bedachte ideeën, maar niet op de originaliteit van de gekozen ideeën.

Uit deze studies kunnen verschillende conclusies getrokken worden (**Hoofdstuk 5**). De resultaten laten ten eerste zien dat productiviteit, het voornaamste doel van een brainstormsessie, niet voldoende is om te komen tot de selectie van creatieve ideeën. Zonder specifieke selectiecriteria kiezen mensen liever voor uitvoerbare, praktische ideeën dan voor creatieve, originele ideeën. Ten tweede laten mijn resultaten zien dat interactieve groepen, ondanks hun lagere productie, niet noodzakelijkerwijs slechtere ideeën kiezen dan nominale groepen. Dit betekent dat het gebruik van interactieve groepen minder rampzalig is dan op basis van eerder brainstormonderzoek geconcludeerd is. Ten derde laten mijn resultaten zien dat de relatie tussen kwantiteit en kwaliteit complexer is dan vaak wordt aangenomen. Diepere verkenning van kennis over een brainstormonderwerp blijkt te leiden tot het bedenken van originelere ideeën. Het bedenken van originelere ideeën leidt echter niet automatisch tot het *kiezen* van originele ideeën.

Er zijn diverse praktische aanbevelingen die uit deze studies kunnen worden afgeleid. Ten eerste is het belangrijk om aandacht te besteden aan het selecteren van ideeën als apart proces. Het kiezen van creatieve ideeën is een vaardigheid op zich, die niet eenvoudigweg bevorderd wordt door te zorgen dat er meer creatieve ideeën voorhanden zijn. Ten tweede, en hiermee samenhangend, is het belangrijk om bij het selecteren van ideeën de juiste selectiecriteria te hanteren, en te zorgen dat deze op de juiste manier geïnterpreteerd worden. Als mensen verondersteld worden criteria te hanteren die ze zelf als tegenstrijdig zien, zal de selectie hier niet door verbeterd worden. Ten derde kan de kwaliteit van bedachte ideeën significant verbeterd worden door een probleem of vraag op de juiste manier te formuleren of te benaderen. De creatiefste ideeën worden bedacht wanneer een probleem zodanig geformuleerd wordt dat diepe verkenning van de relevante domeinkennis aangemoedigd wordt, of zelfs noodzakelijk is. Niettemin bevestigen mijn resultaten dat het bedenken van creatieve ideeën maar één stap in het creatieve proces is: productiviteit is niet genoeg.

## Biography

Eric Fulco Rietzschel was born on September 23rd, 1973, in Delft. He grew up in Voorburg, where he finished secondary school in 1992. In that same year, he entered the University of Amsterdam to study Psychology. In 2000, he graduated cum laude in Psychonomics. From 2001 until 2005, he worked as a PhD student at the University of Amsterdam, in the Department of Work and Organizational Psychology; the research reported in this dissertation was conducted in this period.

Eric currently works as an assistant professor at the Department of Social and Organizational Psychology of the University of Groningen.

**The "Kurt Lewin Institute Dissertation Series" started in 1997. Since 2003 the following dissertations have been published:**

2003-1: Femke Bennenbroek: *Social comparison and coping with radiation therapy. The significance of different dimensions of compariso*n
2003-2: Marlene Roefs: *Public Participation and Perceived (in)justice in South Africa, 1995-2000*
2003-3: Rick van Baaren: *Mimicry: A social perspective*
2003-4: Michaéla Schippers: *Reflexivity in teams*
2003-5: Heleen van Mierlo: *Self-managing teamwork and psychological well-being*
2003-6: Ruben Taris: *Person-Environment Fit: A longitudinal study of the interaction between employees characteristics and work environmental characteristics*
2003-7: Rob Holland: *On the structure and consequences of attitude strength*
2003-8: Camiel Beukeboom: *How mood turns on language*
2003-9: Joost Miedema: *Fairness and the Self*
2004-1: Wolfgang Steinel: *Misleading in social decision-making: A motivational approach*
2004-2: Edwin van Hooft: *Job seeking as planned behavior. In search of group differences*
2004-3: Kirsten Ruys: *The influence of self and similarity on immediate affective impressions of people*
2004-4: Michel Handgraaf: *Self-interest and other-oriented motives in social decision making: An ultimatum bargaining perspective*
2004-5: Sjoerd Goslinga: *Betrokkenheid, participatie en ledenverloop in vakbonden*
2004-6: Nynke Frieswijk: *Frail, but happy: The importance of self-management ability and social comparison among elderly persons*
2004-7: Jaap Ham: *Bridging attribution and spontaneous inferences: Spontaneous and intentional components of dispositional and situational inferences*
2004-8: Mirjam Tazelaar: *When things go wrong: The influence of noise on interactions*
2004-9: Marjolein de Best-Waldhober: *Coördinatie in interdependente situaties: Over de afweging van directe uitkomsten versus uitkomsten op lange termijn*
2004-10: Hinke Groothof: *When others are doing better or worse: Responses from the heart and the head*
2004-11: Laetitia Mulder: *Throwing light on the dark side of sanctions. Sanctioning systems in social dilemmas re-examined*
2004-12: Annebel de Hoogh: *Personality and charismatic leadership: A matter of context*
2004-13: Brigitte ten Brink: *Psychological contract: A useful concept?*
2004-14: Gerben van Kleef: *Emotion in social conflict. The interpersonal effects of emotions in negotiations*
2004-15: Tom Frijns: *Keeping secrets: Quantity, quality and consequences*
2005-1: Elsbeth Steenland: *Stereotype Change: Effects of disconfirming information*
2005-2: Annet de Lange: *What about causality? Examining longitudinal relations between work characteristics and mental health*
2005-3: Reint Jan Renes: *Sustained volunteerism: Justification, motivation and management*
2005-4: Aloyse Augusta Dinsbach: *Socialization in the workplace: A focus on migrant groups*
2005-5: Susanne Peters: *The social psychology of being better off than others*

2005-6: Winny Bakker: *Emigration and well-being: The role of personality and cultural identity in acculturation*
2005-7: Ruud Zaalberg: *The expression of emotion in social situations. The mediating roles of display rules and social motives*
2005-8: Catharine Evers: *Sex differences in anger expressions. The shaping role of social appraisals*
2005-9: Ed Sleebos: *The Consequences of Perceived Intra-group Respect: The Effects of Differential Intra-group Respect on Behavior and Cognition*
2005-10: Jeroen Stouten: *Virtue Summons the Fury: Coordination Rules and Fairness in Social Dilemmas*
2005-11: Irene de Pater: *Doing Things Right or Doing the Right Thing: A New Perspective on the Gender Gap in Career Success*
2005-12: Rob Nelissen: *Guided by Reason, Struck by Emotion: Integrating Motivational & Expectancy - Value Accounts of Behavior*
2005-13: Emely de Vet: *Testing the Transtheoretical Model: Validity and Applicability for Fruit Intake*
2005-14: Floor Rink: *Diversity and Small Group Decision Making: Towards a social identity framework for studying the effects of task-related differences on dyads and groups*
2005-15: Sven Zebel: *Negative Associations: The Role of Identification in Group-Based Guilt*
2005-16: Eric Rietzschel: *From quantity to quality: Cognitive, motivational and social aspects of creative idea generation and selection*