Conceptions of Creativity in Elementary School Mathematical Problem Posing

Benjamin Michael Dickman

Submitted in partial fulfillment of the
requirements for the Degree of Doctor of Philosophy
under the Executive Committee of the
Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2014

ABSTRACT

Conceptions of Creativity in Elementary School Mathematical Problem Posing

Benjamin Michael Dickman

Mathematical problem posing and creativity are important areas within mathematics education, and have been connected by mathematicians, mathematics educators, and creativity theorists. However, the relationship between the two remains unclear, which is complicated by the absence of a formal definition of creativity. For this study, the Consensual Assessment Technique (CAT) was used to investigate different raters' views of posed mathematical problems. The principal investigator recruited judges from three different groups: elementary school mathematics teachers, mathematicians who are professors or professors emeriti of mathematics, and psychologists who have conducted research in mathematics education. These judges were then asked to rate the creativity of mathematical problems posed by the principal investigator, all of which were based on the multiplication table. By using Cronbach's coefficient alpha and the intraclass correlation method, the investigator measured both within-group and among-group agreement for judges' ratings of creativity for the posed problems.

Previous studies using CAT to measure judges' ratings of creativity in areas other than mathematics or mathematics education have generally found high levels of agreement; however, the main finding of this study is that agreement was high only when measured within-group for the psychologists. The study begins with a review of the literature on creativity and on mathematical problem posing, describes the procedure and results, provides points for further consideration, and concludes with implications of the study along with suggested avenues for future research.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Need for the Study

Silver (1994) draws the connection between creativity and problem posing as identified by both mathematicians (e.g., Hadamard, 1945) and creativity theorists (e.g., Getzels & Csikszentmihalyi, 1976). For Silver, the link between creativity and mathematical problem posing exists simply by the latter's presence in a battery of tests for the former. These tests include work by Balka (1974) and Torrance (1966). Silver goes on to comment that "it is reasonable to assume there is some link between [mathematical problem] posing and creativity," but that "the general relationship between [them] is unclear" (p. 21).

Creativity

Creativity as an area of study dates back at least to a lecture given by the American Psychological Association's (APA) president, J.P. Guilford, as his presidential address (Guilford, 1950). In his talk, Guilford called for a concerted effort by the field of psychologists to investigate creativity, and proposed it be carried out by examining common traits among individuals regarded as being *creative*. Over the following half century, many other theorists explored the meaning of 'creative' or 'creativity'. In Treffinger et al (2002) the authors remark that "Treffinger (1996) reviewed and presented more than 100 different definitions [of creativity] from the literature" (p. 5). The number of ways in which creativity is conceived of has now grown to the extent that Sawyer (2012) comments "defining creativity may be one of the most difficult tasks facing the social sciences" (p. 11). To this end, explicitly linking creativity and problem posing is a difficult task.

Mathematical Problem Posing

Brown and Walter's (1990) *The Art of Problem Posing* dedicates an entire chapter to connecting problem posing with problem solving, the latter of which had already occupied an important area of research in mathematics education as inspired several decades earlier by Polya's (1945) *How to solve it*. In addition to laying the foundation for the study of mathematical problem solving, Polya's book also contains numerous instances of problem posing, as evidenced by their presence in several of the heuristics for solving problems: create a simpler problem, create a related problem, create a more general problem, and so forth. More generally, the observation that problem solving involves problem posing can be attributed to Duncker (1945); thus, as is the case with the study of creativity, the study of problem posing dates back to at least the mid-twentieth century.

With regard to mathematics education, the importance of problem posing has been indicated by both mathematicians and mathematics educators (Silver, 1994). Moreover, problem posing features in key mathematical standards documents (e.g., NCTM, 1989; NCTM, 1991) and, according to Kilpatrick (1987), problem formulation should be both a goal *and* means of instruction in the mathematics classroom.

Connecting Creativity and Problem Posing

Given the importance of both problem posing and creativity within mathematics education and the lack of understanding with regard to their connections, there is a clear need to investigate the relationship between the two. Furthermore, considering the variety of definitions provided for the latter, it is necessary to ensure that stakeholders in mathematics education are in agreement as to what can be considered creative, so that efforts to foster creativity in students and educators are not impeded by differences in intended meaning.

Purpose of the Study

The purpose of this study is to understand better the connection between problem posing and creativity by asking relevant judges to assess the creativity of the mathematical problems posed. More precisely, this study concerns how elementary school mathematics teachers, mathematicians, and psychologists working in mathematics education conceive of creativity with regard to mathematical problem posing. Mathematics teachers were chosen for the direct instruction they provide to the students; mathematicians were chosen because they design many of the curricula for young students and because they may be responsible for students' future mathematical education; and psychologists were chosen because they conduct much of the research with regard to mathematical learning among children.

Given the three different groups identified herein, namely, elementary school mathematics teachers, mathematicians, and psychologists who work in mathematics education, the specific questions to be answered by this study are as follows:

1) Is there intergroup agreement among mathematics teachers, mathematicians, and psychologists as to how creativity is conceived of in the context of mathematical problem posing?

2) Is there intragroup agreement among mathematics teachers with regard to how creativity is conceived of in the context of mathematical problem posing?

3) Is there intragroup agreement among mathematicians with regard to how creativity is conceived of in the context of mathematical problem posing?

4) Is there intragroup agreement among psychologists working in mathematics education with regard to how creativity is conceived of in the context of mathematical problem posing?

Procedures of the Study

The study proceeded as outlined in, e.g., Baer and McKool (2009). In particular, the Consensual Assessment Technique (CAT) was used to evaluate the creativity of the mathematics problems, which were posed based on the framework of Brown and Walter (1990) and were formulated using the multiplication table as a starting scenario as suggested in Trivett (1980). These mathematics problems were sent to judges in the three specified categories, namely, mathematics teachers, mathematicians, and psychologists working in mathematics education. Subsequently, the expert judges filled out corresponding forms designed by the principal investigator; the forms are straightforward modifications of those found in Amabile (1996) and Baer (1993) to assess the problems along several dimensions, including judges' own subjective definitions of creativity. Finally, the ratings data were analyzed for intergroup and intragroup agreement among the expert judges using appropriate statistical methods. Specifically, Cronbach's coefficient alpha and the intraclass correlation method were used to determine agreement among-groups and within-groups insofar as how mathematics teachers, mathematicians, and psychologists rate the creativity of the posed problems.

CHAPTER II

LITERATURE REVIEW

Before examining the creativity involved in mathematical problem posing, we begin by taking a step back to consider the study of creativity more generally. The intention in doing so is to present a brief review of the breadth of ways in which creativity is conceived of; in this respect, the diversity of definitions is relevant to the study at hand insofar as it motivates the underlying question of whether or not different stakeholders in mathematics education agree on what constitutes that which is ‘creative’. The study itself ultimately considers creativity as being domain specific, though this feature is still debated within creativity research (e.g., Baer, 1998; Plucker, 1998). Moreover, the perspectives explored are primarily from research and writing originating in the United States; for further background on Western views of creativity from a European standpoint, see intellectual historian Mason’s (2003) *The Value of Creativity*, which traces various conceptions beginning with Genesis in the Old Testament, and moves through Ancient Greece all the way up to work by Nietzsche at the end of the nineteenth century. For a comparison of Eastern and Western conceptions, see, e.g., Hennesey’s (2003) work comparing creative motivators in American and Saudi Arabian culture; the discussion of composing one’s own creative life in the context of women affected by the Iranian revolution (Bateson, 2001); and, with regard to mathematical problem posing, a comparison between high school students in China and the United States (Van Harpen & Sriraman, 2013). More generally, see Lubart (1999; 2010) on cross-cultural perspectives of creativity.

The structure of the literature review is an adaption of the framework in Rhodes (1961) in which one begins by deciding where to *put* the creativity; more precisely, whether the discussion ought to center around a creative person, creative process, or creative product. In each case, such a choice brings with it corresponding frameworks both with regard to the ways in which creativity is defined and the ways in which it is evaluated (Almeida et al, 2008).

Previous studies on the creativity in mathematical problem posing have been organized in other manners; for example, Sriraman (2004) adheres to the approach taken in *The Handbook of Creativity* (Sternberg, 2000) where creativity in a given context is assigned to one or more of the following categories: mystical, pragmatic, psychodynamic, psychometric, cognitive, or social-personality. Similarly, Treffinger and Selby (1993) consider a *COCO model* in which creativity is classified as being produced through the interactions between *Characteristics*, *Operations*, *Context*, and *Outcomes*.

In the subsections to follow, we delve into the literature on creative people and generalize to groups; creative processes and the relation to problem solving; and creative products along with the ways in which they are evaluated. The study here places the creativity in the product – i.e., the mathematical problems posed from the multiplication table – with the result that a deeper discussion of this choice and an associated method of assessment will be necessary. Finally, the domain specificity of the study entails that both mathematical creativity and mathematical problem posing be covered; the former is interwoven into the first three subsections, while the latter concludes the literature review.

Creative People

Guilford (1950) gave his presidential address to the American Psychological Association with a talk entitled *Creativity*. His working definition posits that "creativity refers to the abilities that are most characteristic of creative people. Creative abilities determine whether the individual has the power to exhibit creative behavior to a noteworthy degree" (p. 444). The circular nature of such a definition is not fully resolved, though he remarks that:

> Creative personality is… a matter of those patterns of traits that are characteristic of creative persons. A creative pattern is manifest in creative behavior, which includes such activities as inventing, designing, contriving, composing and planning. People who exhibit these types of behavior to a marked degree are recognized as being creative (p. 444).

The question of what constitutes a creative personality is thereby passed to the question of what constitutes creative behavior, of which Guilford provides five broad examples ("investing, designing, contriving, composing, and planning").

The traits hypothesized by Guilford as characterizing a creative personality include an individual's "sensitivity to problems, ideational fluency, flexibility of a set, ideational novelty, synthesizing ability, analyzing ability, reorganizing or redefining ability [in the sense of Gestalt psychology], span of ideational structure [i.e., complexity/intricacy of an individual's conceptual structure], and evaluating ability" (p. 454). The attempt at classifying individuals in this way is reminiscent of earlier work by Guilford on American military personnel (e.g., Guilford & Lacey, 1947) and followed in his later work as well (e.g., Guilford, 1957).

The first trait mentioned by Guilford, sensitivity to problems, arises in work on *problem finding*, i.e., the ability to describe, formulate, or create problem situations (e.g., Getzels & Csikszentmihalyi, 1976; Runco, 1994). Getzels (1975) quotes Albert Einstein as stating that "The formulation of a problem is often more essential than its solution, which may be merely a matter of mathematical or experimental skill. To raise new questions, new possibilities, to regard old questions from a new angle, requires creative imagination and marks real advances in science" (p. 12). Analogous remarks can be found in Gestalt psychology, where Wertheimer (1959) remarks that "Often in great discoveries the most important thing is that a certain question is found. Envisaging, putting the productive question is often more important, often a greater achievement than solution of a set question" (p. 141). And, in a similar vein, the great mathematician Georg Cantor (1867) wrote in his doctoral dissertation: "In re mathematica ars proponendi pluris facienda est quam solvendi" ("In mathematics the art of asking questions is more valuable than solving problems").

Other traits put forth by Guilford were captured by Torrance's (1968) work analyzing creative characteristics in his development of the Torrance Tests for Creative Thinking (TTCT). This was part of a longitudinal research program to identify creative students early on in their education. The tests themselves primarily evaluate what are referred to as divergent thinking abilities, e.g., *fluency*, *flexibility*, *originality,* and *elaboration*. Fluency refers to the number of relevant ideas produced in a set period of time; flexibility refers to the range of categories corresponding to the ideas generated by a test-taker; originality refers to the generation of ideas not produced by other test-takers; and elaboration refers to the details and development depth of ideas generated. All four divergent thinking constructs can be found in Guilford's (1956) earlier work; respectively, they correspond to ideational fluency, flexibility of a set, ideational novelty, and span of ideational structure. Thus, divergent thinking constructs constitute only a proper subset of the creative traits put forth by Guilford.

It should be noted that divergent thinking measures only one component of an individual's potential for creative thinking (Runco, 1991). Kim (2006) notes that various incarnations of TTCT can be useful in identifying gifted students, and recommends that "considering the multidimensional nature of the creativity concept… at least two measures [to assess creativity] should be used; that is, the TTCT and another indicator (e.g., products, performance, rating scales, or recommendations)" (p. 11). Runco (1993) also comments that while the "evolution of the divergent thinking paradigm is very important for those interested in giftedness… recent research indicates that it is no longer sufficient to look to general divergent thinking when attempting to identify gifted children" (p. 21). He goes on to suggest that, in the case of children, educators should specifically complement divergent thinking tests by also looking at problem finding and problem defining capabilities.

In particular, although the TTCT was initially administered as a method of evaluating creative thinking, it is an instrument designed specifically to measure divergent thinking; in this respect, Runco (2010) remarks that divergent thinking "is not a synonym for creativity but is useful for research on creative potential" (p. 414). As a key researcher on divergent thinking, Runco reiterates many times the point that divergent thinking and creative thinking are not one and the same; see, for example, *Commentary: Divergent thinking is not synonymous with creativity* (2008) and a re-statement elsewhere by Runco (1993) that, while it is difficult to measure or define creativity, "divergent thinking tests are often used, though… they really just *estimate* the potential for creative thought" (p. 16).

Others have criticized Torrance's test on statistical grounds; for example, Almeida et al (2008) carried out a factor analysis on three TTCT administrations, and found its construct validity questionable insofar as "cognitive functions supposedly present on a subject's performance are not so strong as to explain the variance in scores" (p. 53). Note that the testing locations in this study, Spain and Portugal, may make the results less relevant for the United States. Davis (1997) comments that, in spite of criticism concerning the validity of TTCT, the data seem reasonable; he goes on to provide several articles summarizing pros and cons of the tests (Callahan, 1991; Chase, 1985; Treffinger, 1985; Torrance, 1979, 1988, 1990).

One response to the trait-based theories designed to investigate creative people can be found in the work of Gruber (e.g., Gruber, 1988; Gruber & Davis, 1988; Gruber & Wallace, 1999). Gruber conceives of an individual as an evolving system that is composed of loosely coupled subsystems of *knowledge, affect,* and *purpose,* and considers how these different parts evolve and interact with one another over time. He also considers the individual's *network of enterprises,* which refers to the different projects in which one is engaged. A brief summary of Gruber's case study approach to researching creative people can be found in Brower (2003), where the author demonstrates how key features of this methodology can be used in an investigation of Vincent van Gogh, to which Brower adds a specific focus on the role of creative repetition in van Gogh's artistic development.

Gruber's approach to studying creativity traces directly to his own lengthy case study on Charles Darwin's development of his theory of evolution (Gruber & Barrett, 1974). As the method of study was further refined and applied to other people widely considered to be creative, e.g., Benjamin Franklin and John Locke, various foci emerged with regard to the ways in which the creativity theorist can try to investigate the evolution of an individual's creativity (Gruber & Davis, 1988). Besides Brower's (2003) focus on creative repetition, there is a focus on metaphors, thought-forms, and, subsequently, a list of nine facets for consideration in performing such a case study (Gruber & Wallace, 1999). These facets are: uniqueness, epitome, systems of belief, modality of thought, multiple timescales, purposeful work and network of enterprises, problem solving, contextual frames, and values.

An example of Gruber's method applied within the domain of mathematics can be found in Dickman (2013). Paul Cohen's development of the mathematical technique of *forcing* is examined by investigating his modality of thought – i.e., Cohen's tendency to think in terms of set theoretical models – and a system of beliefs with regard to decision procedures where he wishes to resolve mathematical problems by first working with the simplest ones, and then building up inductively to tackle those of greater complexity. The aforementioned facets of

Cohen's creative work combine and evolve along with a sense of purpose honed by his exposure to the priority method, a solution to a different problem appearing five years before Cohen's seminal paper, which served as one of the key precursors to forcing (Kunen, 1980; Moore, 1988).

Cohen (2002) writes up his own backstory on the development of forcing, and, in this respect, provides an example of a highly creative individual giving a first-hand account of his own work. A number of eminent individuals, ranging from physicist Richard Feynman to musician Frank Zappa, provide short accounts of their own creative work in Barron et al (1997). With regard to mathematicians, in particular, Poincare (1982) and Hadamard (1945) provide their own personal accounts; more on these sources can be found in the next section.

There are at least two dangers in relying primarily on individuals' own accounts of their creativity. One is that *post hoc* recollections of creative inspiration are often less accurate (Gruber, 1981). For example, Cohen (2002) makes no mention of the priority method, despite its later influence on his development of forcing (Dickman, 2013; Cohen, 2008). A second potential pitfall is that telling one's own story can constitute a form of creativity in itself (Bateson, 2001). More precisely, Bateson notes that in telling a story about one's own life, certain details will necessarily be omitted, and that individuals may choose to frame their personal narratives with an emphasis on either continuities or discontinuities in their experiences. She notes that this framing can have an effect on both the storyteller as well as the audience (Bateson, 2004). A particular example provided in this context relates to how parents relay stories to their children, where an overemphasis on continuity can gave a discouraging false impression of the reality of life's twists and turns. The salient point with regard to creativity is that stories of creative individuals can be presented in a way so as to make them seem more natural in retrospect than they were when lived. Just as Gruber works to demystify the creativity of Charles Darwin, other theorists' investigations can do the same for others, e.g., Weisberg (1988; 2004; 2011) on Watson and Crick, Picasso, and Frank Lloyd Wright, respectively.

Sawyer and DeZutter (2009) move beyond discussing creative people on an individual basis, and talk about the *collaborative emergence* of creativity in group settings. The scenarios they focus on are those with four primary characteristics: first, the activity does not have a predictable outcome; second, there is moment-to-moment contingency, meaning that each action depends on its immediate predecessor; third, the interactional effects of a given action are subject to change with respect to subsequent actions by other participants; and, fourth, individual contributions are equally distributed within the collaborative process.

The primary method for studying collaborative emergence of creativity within groups is *interaction analysis*, which consists of videotaping group members' collaboration over time, coding the videos for individual and group contributions and the changes they undergo, and asking group members to provide their own input insofar as how these interactions are experienced personally (Sawyer & DeZutter, 2009). Improvisation activities provide scenarios that are both amenable to interaction analysis and pedagogically relevant (e.g., Sawyer, 2011; Lobman & Lundquist, 2007). Sawyer notes further that, in the context of education, the tensions between structure and improvisation can lead to difficulties in the classroom and curriculum.

Problems may arise as creativity emerges among students in the classroom. Cropley (2010) distinguishes between orthodox and radical creativity, where the latter can interfere with both individual students' school learning as well as the learning of their peers. More generally, he notes that creativity emerging in schools can lead to several potential issues: creativity brings with it an inherent uncertainty for both students and their parents; it can lead to the questioning of working hard to acquire skills and knowledge; student creativity can threaten teachers' status and authority; and it can lead to a weakening of teachers' self-image. Brown and Walter (1990) observe an analogous shift in authority when students are allowed to create their own mathematical problems; indeed, each of the five potential issues above can be applied to classrooms that incorporate problem posing into the curriculum as well.

Creative Processes

Prior to the shift towards discussing the construct of *creativity* in the mid-twentieth century, much of the work that would be later drawn upon is found in more general discussions of *thinking*. An early, seminal piece is Wallas' (1926) *The Art of Thought*, in which he puts forth a four-stage model of problem solving: preparation, incubation, illumination, and verification. The *preparation* involves background content-knowledge, which suggests again domain specificity, as well as conscious work on the problem at hand. *Incubation* refers to an interstitial lull during which the solver spends time away from consciously working on the problem. *Illumination*, sometimes known as *insight*, refers to the solution's emergence into an individual's conscious thinking. Lastly, *verification* denotes the solver's checking, assessment, and implementation of a solution.

Wallas' four-stage model is frequently incorporated or adapted in works on the creative thought process. For example, Torrance and Safter's (1990) *incubation model* is summarized in Starko (2004) as consisting of three components: insight, intuition, and revelation. Starko discusses this model as beginning with the heightening of anticipation, during which individuals prepare to connect that which they are expected to learn with their own lives in some meaningful way; proceeds with the deepening of expectations, when the individual both processes new information and consciously works on problems that emerged during the previous stage; and, after an incubation period, concludes with individuals going beyond: for example, using their creative output to solve problems in the real world.

Mathematics provides an example of Wallas' general model in the autobiographical recollection of Poincare (1982). After consciously thinking about Fuschian functions for some stretch of time, it is only subsequent to an incubation period – during which his attentions are directed elsewhere – that Poincare steps onto an omnibus and suddenly has the epiphany that allows him to forge onward in his mathematical work. Quoting the English translation:

> Most striking at first is this appearance of sudden illumination, a manifest sign of long, unconscious prior work. The role of this unconscious work in mathematical invention appears to me incontestable, and traces of it would be found in other cases where it is less evident. Often when one works at a hard question, nothing good is accomplished at the first attack. Then one takes a rest, longer or shorter, and sits down anew to the work. During the first half-hour, as before, nothing is found, and then all of a sudden the decisive idea presents itself to the mind (p. 389).

Other mathematicians report similar experiences. Hadamard (1945) writes of Poincare's discovery as well as his own work. For example, he provides his own recollection of a mathematical insight shortly after arising:

> One phenomenon is certain and I can vouch for its absolute certainty: the sudden and immediate appearance of a solution at the very moment of sudden awakening. On being very abruptly awakened by an external noise, a solution long searched for appeared to me at once without the slightest instant of reflection on my part – the fact was remarkable enough to have struck me unforgettably – and in a quite different direction from any of those which I had previously tried to follow (p. 8).

Hadamard also mentions a quotation of Gauss:

> Finally, two days ago I succeeded, not on account of my painful efforts, but by the grace of God. Like a sudden flash of lightning, the riddle happened to be solved. I myself cannot say what was the conducting thread which connected what I previously knew with what made my success possible (p. 15).

Cohen (2002) similarly recalls his own work on developing forcing as having occurred while driving near the Grand Canyon with his wife; of his epiphany at that moment he writes:

> There are certainly moments in any mathematical discovery when the resolution of a problem takes place at such a subconscious level that, in retrospect, it seems impossible to dissect it and explains its origin. Rather, the entire idea presents itself at once, often perhaps in a vague form, but gradually become more precise (p. 1092).

Wertheimer (1953) provides his own account of productive thought as constituting “one consistent line of thinking” rather than a “sum of aggregated, piecemeal operations” (p. 42). For Wertheimer, the thought process commences with a desire to understand how different objects relate to one another, and continues with such operations as *grouping*, *reorganization,* and *structurization*. With regard to successful problem solving, Wertheimer writes:

> I would say that the essential features in genuine solving are as follows: not to be bound, blinded by habits; not merely to repeat slavishly what one has been taught; not to proceed in a mechanized state of mind, in a piecemeal attitude, with piecemeal attention, by piecemeal operations; *but* to look at the situation freely, open-mindedly, viewing the whole, trying to discover, to realize how the problem and the situation are related, trying to penetrate, to realize and to trace out the inner relation between form and task; in the finest cases getting at the roots of the situation, illuminating and making transparent essential structural features… in spite of the difficulties (pp. 120-121).

Thus, Wertheimer advocates for a holistic approach to problem solving, where one can engage in divergent thinking when necessary, and where “the finest cases” of productive thought are characterized not only by overcoming structural complexity, but by understanding previously opaque structural features in such a way so as to make them entirely transparent.

Cognitive psychologist Stokes (2005) discusses what constitutes creativity and how it can be developed in her book *Creativity from Constraints: The Psychology of Breakthrough*, in which she argues that the use of constraints can greatly enhance the creative process. She defines creativity as a commendation given to responses that are new and generative or influential as well as appropriate or useful (Stokes, 2006; Stokes, 2010). *Generative* means that the output leads to the generation of other ideas and things; *influential* means others’ responses to similar work change as a result of the output; and *appropriate* or *useful* means that the output solves some sort of problem. Given these definitions as a springboard, Stokes proceeds to discuss what she calls *the creativity problem*, which has a twofold characterization. First, in keeping with the creativity literature from cognitive psychology more generally, the problem space is ill-structured, in the sense that it is incompletely specified and does not have a clear goal criterion. Second, the problem's solution requires the specification and subsequent strategic use of pairs of constraints; in particular, these are *precluding constraints* and *promoting constraints*, where the former are constraints that limit one's search among known responses, and the latter are constraints that promote the search for novel solutions.

The four main types of constraints discussed are domain constraints, talent constraints, cognitive constraints, and variability constraints (Stokes, 2005). Lastly, Stokes emphasizes the importance of what she calls *first choruses*, which refers to a mastered domain upon which future work is improvised. Note the inherent domain specificity insofar as each domain has its own set of performance criteria; namely, goal constraints, subject or content constraints, and task constraints. Drawing from one's first chorus is an essential part of the creative process, though the domain itself may be embellished upon or have certain aspects emphasized, and may even lead a creator to draw from outside domains as well. Insofar as this study is concerned, the salient point is that the creative person is the one who has sufficient domain-expertise so as to impose the proper constraints in order to produce something creative; for example, with respect to problem posing, a mathematics teacher might restrict to a single mathematical scenario (such as the multiplication table) when formulating questions.

Weisberg (2006) believes that creative thinking is an example of ordinary thinking, where the latter consists of such cognitive components as: remembering, imagining, planning, and deciding, More generally, he characterizes ordinary thinking by remarking that our thoughts have structure, are continuous with the past, are sensitive to the environment, and are directed by *top-down processing*. By top-down processing, or concept-driven processing, Weisberg means "the use of knowledge and expectations in cognitive functioning" (p. 114). As we use our senses to analyze input from the environment, this perceptual processing is complemented by our knowledge and expectations about the world; together, these lead to problem solving and creative thinking.

More generally, Weisberg concludes his book by identifying creativity with problem solving, and writes that "it seems reasonable to adopt as a working assumption the premise that creative thinking is an example of problem solving" (p. 581). He addresses the earlier study on problem finding by Getzels and Csikszentmihalyi by reclassifying it as an example of participants who "were actually faced with *two* problems" (p. 140). In particular, the first problem was to choose and arrange the objects to be painted, and the second problem was to paint them. With regard to mathematical problem posing with a multiplication table, one might deconstruct this in a similar manner to Weisberg by framing it as, first, choosing a mathematical scenario (the multiplication table), and, second, formulating questions related to the chosen scenario.

Creative Products

In light of the creative trichotomy presented, products are certainly the easiest to assess. Davis (1997) cites Borland in Colangelo and Davis (1997) when noting that "evaluating actual products reflects the way talent and capability are recognized in the real world" (p. 274). He also discusses the use of a particular technique of product assessment, known as the Consensual Assessment Technique (CAT), that originates in work by Amabile and Hennessey (e.g., Amabile, 1983; Hennessey, 1994; Hennessey & Amabile, 1999). This methodology has been used in numerous tasks, ranging from dramatic performance to musical compositions to personal narratives; see Baer and McKool (2009) for a host of citations for these and other tasks.

The Consensual Assessment Technique is summarized in, e.g., Baer and McKool (2009) as consisting of two basic components: first, asking subjects to create something; second, asking experts in the relevant domain to evaluate the creativity of the items. The experts' ratings are then assessed using interrater reliability tests. Though CAT provides implicit support of domain specificity insofar as expert selection is concerned, the interrater reliabilities, measured most often using Cronbach's coefficient alpha, generally range from 0.70 to 0.90 regardless of the domain considered. With regard to those selected to assess creative products, Baer and McKool note that in Amabile (1996) "the average number of expert judges reported… was just over 10, with a low of 2 and a high of 40" (p. 5).

Amabile (1983; 1996) developed her theory of creativity based on the observation that all evaluations of creative products involve some sort of judgment. Even the TTCT, which measures divergent thinking as an estimate of creative potential, is scored by judges who subjectively assign ratings. With regard to her own conception of creativity, Amabile (2012) writes:

> Two important assumptions underlie the theory. First, there is a continuum from low, ordinary levels of creativity found in everyday life to the highest levels of creativity found in historically significant inventions, performances, scientific discoveries, and works of art. The second, related underlying assumption is that there are degrees of creativity in the work of any single individual, even within one domain. The level of creativity that a person produces at any given point in time is a function of the creativity components operating, at that time, within and around the person.

Everyday creativity is sometimes referred to as 'little-c' creativity, whereas the higher level of eminent creativity is sometimes referred to as 'Big-C' creativity (e.g., Kaufman & Sternberg, 2006). Outside of Amabile's work, there is a more recent push by some researchers toward 'mini-c' creativity, which relies not on interpersonal judgments, but rather on *intra*personal judgment (e.g., Beghetto & Kaufman, 2007).

Amabile developed the Consensual Assessment Technique as a way of operationalizing the necessarily subjective judgments of creativity into a working definition. Summarized succinctly by Hennessey et al (2011): "a product or response is creative to the extent that appropriate observers agree it is creative. Appropriate observers are those familiar with the domain in which the product was created or the response articulated" (p. 255). The authors follow by discussing the procedural requirements of CAT: choosing judges with domain-relevant expertise; having judges assess products independently; and having the judges rate products relative to one another. In this last respect, Hennesey et al note that the relative ratings are "important because, for most studies, the levels of creativity produced by the 'ordinary' subjects who participate will be very low in comparison with greatest works ever produced in the domain" (p. 256). Lastly, the authors suggest presenting the products in a randomized order to each judge, noting that "if all judgments are made in the same order by all raters, high levels of agreement might reflect methodological artifacts" (p. 256). Subsequently, the ratings can be analyzed using statistical approaches to measure internal consistency, such as Cronbach's coefficient alpha or the intraclass correlation method.

Whereas the Consensual Assessment Technique provides a statistical method of evaluating the creativity of a product at the level of everyday creativity, Csikszentmihalyi's (1999) *systems perspective* allows for a more theoretical approach to examining how products are deemed eminently creative. He distinguishes between those who do and do not include public recognition as a part of what makes something creative, tracing the latter viewpoint back to work by Maslow (1963). Csikszentmihalyi is of the former opinion, remarking that "if creativity is to retain a useful meaning, it must refer to a process that results in an idea of product that is recognized and adopted by others" (p. 314). In this respect, he agrees with Amabile that creativity is, fundamentally, a judgment. He continues:

> In practice, creativity research has always recognized this fact. Every creativity test, whether it involves responding to divergent-thinking tasks or whether it asks children to produce stories or designs with colored tiles, is assessed by judges or raters who weigh the originality of the responses. The underlying assumption is that an objective quality called "creativity" is revealed in the products, and that judges and raters can recognize it. But we know that expert judges do not possess an external, objective standard by which to evaluate "creative" responses. Their judgments rely on past experience, training, cultural biases, current trends, personal values, idiosyncratic preferences. Thus, whether an idea or product is creative or not does not depend on its own qualities, but on the effect it is able to produce in others who are exposed to it (p. 314).

Based on these assumptions, his *systems view of creativity* consists of three interacting components – the individual, the field, and the domain – where a product is considered creative precisely when it effects long-term change in its corresponding domain. The field refers to the gatekeepers or individuals of influence in a domain, and are the ones who ultimately decide what is incorporated therein. For example, in the domain of mathematics, an individual mathematician might prove a theorem or derive a new result, but it cannot be considered creative (in Csikszentmihalyi's framework) unless the corresponding field (e.g., notable mathematicians, editors of mathematical journals, authors of mathematical textbooks) deem it so. In this respect he is at odds with Amabile, who applies the word 'creativity' to products that span everyday creativity to eminent creativity; Csikszentmihalyi reserves the word only for the eminent.

The notion that a product can only be considered creative if others take note of it may be somewhat troubling. With regard to this point, Csikszentmihalyi remarks briefly about Vincent van Gogh, whose paintings are regarded as creative today, but were not during his own life.

> But would he have been creative anyway, even if we didn't know it? In my opinion, such a question is too metaphysical to be considered part of a scientific approach. If the question is unanswerable in principle, why ask it? The better strategy is to recognize that

> in the sciences as in the arts, creativity is as much the result of changing standards and new criteria of assessment, as it is of novel individual achievements (p. 321).

Thus, the possible discomfort resulting from the systems view of creativity is not resolved insofar as it is merely dismissed.

Citing the discussion of cultural units of imitation in Dawkins (1976), Csikszentmihalyi suggests that “it is useful to think about creativity as involving a change in memes” (p. 316). He also notes that the size of a field will vary in accordance with the domain to which it corresponds, writing: “When the domain is arcane and highly codified, like Assyriology or molecular biology, then the decision as to which new meme is worth accepting will be made by a relatively small field” (p. 326). Similarly, Csikszentmihalyi remarks that “in physics, the opinion of a very small number of leading university professors was enough to certify that Einstein’s ideas were creative,” and that “hundreds of millions of people accepted the judgment of this tiny field and marveled at Einstein’s creativity without understanding what it was all about” (p. 315). Thus, even in large domains such as physics, the number of members in the corresponding field need not have a large lower bound.

Mathematical Problem Posing

When providing background information on mathematical problem posing, a bit of history surrounding mathematical problem solving ought to be included as well, for each one begets the other. It is clear that for a problem to be solved, it must first be, in some form, posed. The literature on mathematical problem solving dates back at least to Polya’s (1945) *How to solve it*, in which he provides a general method for dealing with mathematical questions. Polya’s fundamental four steps are to understand the problem, decide on a strategy, carry out the strategy, and then look back. Examples of the last step include questions such as: *Could the* *problem have been solved in a different way? Were all the assumptions necessary? Could the solution be seen at a glance?*

The main content of Polya's opus is a long list of *heuristics*, or strategies, for broaching a problem when one does not know at the outset which method or methods to use. These heuristics include asking oneself questions such as: *Do I know a related problem? Do I know a simpler problem? Can I generalize the problem? Can I work backwards? Can I draw a relevant picture?* In each case, the solver is attempting to make progress by posing new, hopefully more tractable problems. By asking and answering these types of questions, Polya believes the solver will be able to decide on a successful strategy to resolve the initial problem.

One weakness of Polya's book is the extent to which students of mathematics differ in their understanding of how to use the heuristics. For example, Silver (1979; Silver & Smith, 1980) examined solvers' conceptions of what constitutes a *related problem*, and found that mathematicians consistently grouped together problems that were alike with regard to underlying mathematical structure, whereas a significant portion of students grouped problems based solely on superficial (i.e., non-mathematical) features. Schoenfeld and Herrmann (1982) came to similar conclusions in looking at grouping by relevance; Schoenfeld (1979) also noted the difficulty entailed in trying to teach heuristics directly to students of mathematics, particularly with regard to whether students even use a problem-solving technique that has been mastered.

A more complete framework for what is necessary to be a successful mathematical problem solver can be found in Schoenfeld's (1985) book *Mathematical Problem Solving*. Besides *heuristics*, which is what Polya's work concerned itself with, Schoenfeld's four part framework also puts forth *resources*, *beliefs* and *belief systems*, and *control* and *metacognition* as integral features. Resources refers to the body of knowledge that a solver can draw upon when trying to figure out a problem; beliefs refer to a mathematical worldview, and include, for example, both epistemological beliefs and the solver's affect in the problem solving process; and control relates to resource management and knowing when to pursue or abandon a particular avenue of solution. The metacognitive features of control include occasional self-checks on why an approach is being used; Schoenfeld (1987) notes that the difference in self-checks is one aspect that distinguishes experts from novices in mathematical problem solving.

Silver (1994) followed his earlier work on problem solving with an article entitled *On Mathematical Problem Posing*. Both Silver (1994) and Kilpatrick (1987) trace the observation that problem posing plays a role in problem solving back to Duncker (1945). In discussing the role of problem posing in the classroom and in mathematics curricula, Silver comments that "problem posing has been identified by some distinguished leaders in mathematics and mathematics education as an important aspect of mathematics education… and problem posing has recently begun to receive increased attention in the literature on curricular and pedagogical innovation in mathematics education" (p. 19). He continues by pointing to standards documents released by the National Council for Teachers of Mathematics (NCTM, 1989; NCTM, 1991) and cites Kilpatrick (1987) as arguing for problem formulation to be both a goal *and* means of instruction in the mathematics classroom.

Silver (1994) also includes a section called *Problem posing as a feature of creative activity or exceptional mathematical ability*, in which he remarks that the "apparent link between posing and creativity is clear from the fact that posing tasks have been included in tests designed to identify creative individuals" (p. 20). In addition to the problem finding tasks of Getzels and Csikszentmihalyi, and the Torrance Tests for Creative Thinking, Silver notes that Balka (1974) "asked subjects to pose mathematical problems that could be answered on the basis of information provided in a set of stories about real-world situations" (p. 20). As with the TTCT, Balka's test was evaluated in terms of divergent thinking constructs (i.e., fluency, flexibility, and originality). Despite the evidence Silver draws from these alleged creativity tests, he remarks that "the general relationship between creativity and problem posing is unclear" (p. 21).

Later work by Silver and Leung (1997) includes the development of a Test of Arithmetic Problem Posing (TAPP) for prospective elementary school teachers. The TAPP is based on Leung's (1993) dissertation, completed under Silver, and consists of four items that were modified from earlier work on problem finding. With regard to creativity, the authors write that "despite the fact that the TAPP test items were adapted from similar tasks used by Getzels and Jackson to identify creative individuals, the results of this investigation suggested essentially no influence of creativity (as represented by the overall score on TTCT-V) on the quality or complexity of the problems posed (as measured by TAPP)" (p. 20). The TTCT-V refers to a later incarnation of the Torrance Tests for Creative Thinking; in particular, Silver and Leung found that the caliber of problem posing was not affected by divergent thinking scores. The authors suggest that one possible reason is that prospective elementary school teachers generally scored low when assessed using these instruments: "The absence of individuals who scored absolutely high on the TTCT-V may have restricted the ability of the analyses conducted in this investigation to detect the influence of creativity on mathematical problem posing" (p. 20). Note the assumption that the TTCT-V has somehow measured creativity, whereas the earlier discussion indicated these tests measure, at most, an estimate of a potential for creativity.

Leung (Shuk-kwan, 1997) remarks that the TAPP is only appropriate for arithmetic problems, and later developed her Test on General Problem Posing (TGPP). Using this instrument, ninety-six elementary school children in Taiwan completed eighteen TGPP test items, which were divided into three categories: text, answer, and picture. The main finding was that there exists a general problem posing competence, insofar as "no proficiency in any of these three task environments [text, answer, picture] is specific and not general to the other two task environments" (p. 83). Though the author found that some students posed original problems, no significant findings are reported with regard to creativity.

Besides Taiwan, other studies on problem posing have been conducted in Mainland China. For example, Van Harpen and Sriraman (2013) selected high school students from three locations: Normal, Illinois, a Midwestern town in the United States; Shanghai, a large city in China; and Jiaozhou, a small city in China. The number of participants from the three locations were, respectively, 30, 44, and 55; all were chosen as advanced students of mathematics. Participants were given a mathematics content test based on the National Assessment of Educational Progress (NAEP) 12th grade Mathematics Assessment, as well as a problem posing test based on earlier work of Stoyanova and Ellerton (1996). The latter test consists of three types of situations: free, semi-structured, and structured. Data collected were coded, sorted into ten different categories, and then analyzed based on the divergent thinking constructs of flexibility, fluency, and originality. The authors found that those who were capable at scoring high on routine mathematical tests were not necessarily capable problem posers; however, students from Jiaozhou were able to score significantly higher than the US and Shanghai groups on both the mathematical content test and the problem posing test.

The authors of the previous study note that their findings contradict the results in another problem posing study comparing United States and Chinese students (Cai & Hwang, 2002). In this latter study, the authors recruited 98 US students and 155 Chinese students, all in the sixth grade, to participate. Students were asked to complete a problem solving and problem posing test; the latter was analyzed according to the three categories of extension problems, non-extension problems, and other, based on earlier general work on problem posing (Simon, 1979) as well as research specific to mathematical problem posing (Cai, 1998). Sriraman and Van Harpen (2013) remark that the Cai and Hwang study differed from theirs insofar as the latter concludes that "US students performed as well as or better than… Chinese students in problem-posing tasks" (p. 217). However, Cai and Hwang are hesitant to generalize too much from their own findings, stating that "the data from this study are insufficient to make strong conclusions about" the nature of problem posing and its connection with problem solving among US and Chinese students (p. 420).

Outside of assessments of problem posing abilities and the connection to creativity, there remains the question of how to help students (including prospective teachers) to become more effective problem posers. Brown and Walter's (1990) *The Art of Problem Posing* provides a general approach for problem generation. In particular, their strategy consists of choosing a mathematical scenario, listing relevant attributes, and then a stage entitled "What-If-Not-ing," in which the problem poser varies the attributes (p. 61). Example scenarios suggested by Brown and Walter include the Pythagorean Theorem, combinatorics problems concerning handshakes, isosceles triangles, and the use of GeoBoards. For example, a student might note a GeoBoard is square (attribute listing), modify to a circular Geoboard (What-If-Not), and consider the number of nails found in a circular GeoBoard of a particular size (problem posing). The student might elect to list further attributes of the circular GeoBoard and vary them as well; through this process of *cycling* between attribute listing and What-If-Not-ing, problem posers can formulate a diverse set of mathematical questions.

Brown and Walter (2014) subsequently published *Problem Posing: Reflections and Applications*, in which they assemble a compendium of articles by mathematics educators who carried out activities based on their earlier book in a variety of mathematical scenarios. Despite the richness of some of these lessons, there is a marked absence of one of the more common tools found in elementary school mathematics classrooms: the multiplication table. Though the multiplication table is used a source of problems elsewhere (e.g., Abramovich, 2007), it has not been exploited to the extent called for by Trivett (1980), who writes of the table as a “veritable repository of mathematical relationships” (p. 21). From this perspective, the multiplication table is a viable starting point for problem posing. In particular, the table can be used to pose a variety of mathematical problems, which can then be assessed for their subjective creativity by judges who are familiar with the corresponding domain. In keeping with the earlier discussion of the Consensual Assessment Technique, these ratings can then be compared so as to evaluate the extent to which experts agree with one another as to what constitutes creativity in this area of mathematical problem posing.

CHAPTER III

PROCEDURES

The study was completed in four steps. First, the investigator posed twenty-five mathematics problems based on the multiplication table. In the second step, these mathematics problems were sent to expert judges in one of three categories: mathematics teachers, mathematicians, and psychologists. In the third step, the expert judges evaluated all twenty-five items on five different dimensions: creativity, pedagogical appropriateness, liking, clarity, and originality of idea. Finally, the ratings data were analyzed for intragroup and intergroup agreement among the expert judges using appropriate statistical methods. Each of the four steps is outlined below.

Posing Problems with the Multiplication Table

The principal investigator adhered to the suggestion of Trivett (1980) who writes:

> The recommendation here is that the multiplication table should be viewed, apparently for the first time by most people, as a dynamic synergetic combination of patterns, a veritable repository of mathematical relationships waiting as it were to gush forth from kindergarten through the secondary grades (p. 21).

In particular, the principal investigator utilized problem posing techniques based on the methods described in Brown and Walter (1990) to produce the twenty-five problems with the multiplication table as a starting mathematical scenario. The full list of problems is in Appendix B. The problems were written up without solutions; in some cases, there is a unique solution (e.g., Problem 13) while other problems have multiple solutions (e.g., Problem 25). The absence of solutions is due to the focus on problem posing, rather than problem solving.

For explicit examples of the problem posing procedure, consider the following: One attribute of the multiplication table is that, since all entries are natural numbers, each can be classified as either even or odd. A question based directly on this attribute can be found in Problem 2: "How many of the entries in a 10x10 multiplication table are odd?"

Another attribute of the multiplication table is that it only covers the multiplication of positive numbers. A question based on the What-If-Not technique as applied to this attribute can be found in Problem 16: "How would you extend the 10x10 multiplication table to cover the negative numbers?"

Finally, an attribute of the multiplication table is that it has finitely many entries. A possible question based on the What-If-Not-ing technique would be: "What happens when you consider the multiplication table rows and columns both extended indefinitely?" The resulting extended multiplication table has certain interesting attributes; for example, the prime numbers can be defined as the numbers which appear exactly twice. (Note that this is *not* the case for the standard 10x10 table, where, for example, 15 appears exactly twice: once as 3*5 and once as 5*3. But 15 is not prime, as evident by the given decompositions.) By cycling from attribute listing to What-If-Not-ing to attribute listing again, one can pose Problem 9: "Using a multiplication table, explain what it means for a number to be 'prime.'"

Expert Judges

The judges were selected based on recommendations from the Mathematics Education and Developmental Psychology graduate programs at Teachers College Columbia University. Participants are considered *mathematics teachers* if they have taught at least one school year in the United States between the third and sixth grade, during which they were responsible for the mathematics portion and either covered the multiplication table explicitly or taught students who had covered it in previous years. Participants are considered *mathematicians* if they hold a doctoral degree in mathematics and are professors or professors emeriti of mathematics.

Participants are considered *psychologists* if they hold a doctorate in psychology or a related field (e.g., educational psychology, developmental psychology) and work as professors or professors emeriti who conduct research in mathematics education.

The selection of judges is in accordance with an earlier study by Amabile (1983) in which the expert judges were art teachers, artists, and psychologists. For the study at hand, the corresponding choices of mathematics teachers (n = 5), mathematicians (n = 6), and psychologists (n = 9), respectively, are based upon the domain change from art.

All participants were recruited by electronic mail correspondence. After obtaining recommendations for raters through the aforementioned graduate programs, the principal investigator sent out emails entitled *Multiplication Table Problems*. In keeping with the rules set out by the Institutional Review Board (IRB), all potential participants were told that they could decline or withdraw subsequently. Recipients who indicated interest or potential interest in participating received a follow-up email. Those who asked specific questions about how to rate were, in general, instructed to "do their best." Two more follow-up emails were later sent to those who indicated interest in participating but whose forms had not yet been received; the language of all follow-up emails was essentially the same, and is reproduced in Appendix E. In keeping with the methodology of CAT, all emails sent to participants indicated a source for the products to be rated; more precisely, that the problems for analysis were created by "prospective mathematics teachers."

The twenty participants for whom data were analyzed signed and returned the informed consent forms that were approved by the IRB (Protocol #14-178). The IRB approval form can be found in Appendix D.

Ratings and Data Analysis

The expert judges were presented with all twenty-five multiplication table problems and asked to rate them relative to one another using Likert scales from 1 (least) to 5 (most) based on the following dimensions: creativity, pedagogical appropriateness, liking, clarity, and originality of idea. These dimension choices and their descriptions are straightforward adaptations of earlier work by Amabile (1983) and Baer (1993). For creativity, in particular, the working definition provided was: "The degree to which the math problem is creative, using your own subjective definition of creativity."

After data were collected, the ratings were analyzed using the Consensual Assessment Technique. As stated by Baer and McKool (2009) the "inter-rater reliability using the Consensual Assessment Technique is typically measured using Cronbach's coefficient alpha… or the intraclass correlation method" (p. 5). In this study, Cronbach's coefficient alpha and the intraclass correlation method (ICC) were both used to analyze the data; all methods were executed in the statistical software package *R*.

With regard to Cronbach's coefficient alpha, a score greater than or equal to 0.70 is considered a high measure of internal agreement (Baer & McKool, 2009). This statistical test is carried out several times: it is applied to the full matrix of ratings in order to measure the intergroup agreement for the twenty judges, and to each of the individual groups of judges in order to measure intragroup agreement. For the five hundred creativity ratings, only one judge, P5, omitted a problem: the judge indicated that she could not understand Problem 22 well enough to rate it. Rather than removing this problem entirely, the principal investigator chose to mark the creativity for this problem as a 1. All other ratings are precisely as reported.

In addition to Cronbach's alpha, ICC was carried out to evaluate intragroup agreement for the mathematics teachers, mathematicians, and the psychologists who work in mathematics education. Baer and McKool (2009) remark that "these methods generally yield similar inter-rater reliability estimates" (p. 5). Indeed, the statistical conclusions were similar for both Cronbach's alpha and ICC with regard to evaluating intragroup agreement.

CHAPTER IV

RESULTS

The principal method of statistical analysis used for the Consensual Assessment Technique relies on Cronbach's coefficient alpha to evaluate agreement among judges. The intraclass correlation method (ICC) is another approach to assessing agreement, and was carried out for each of the individual groups.

In all sections and subsections to follow, 'T' denotes 'mathematics teacher', 'M' denotes 'mathematician', and 'P' denotes 'psychologist'.

Intergroup Agreement

The question addressed here is whether or not there exists intergroup agreement among mathematics teachers, mathematicians, and psychologists as to how creativity is conceived of in the context of mathematical problem posing.

Cronbach's coefficient alpha for the entire group of judges ($n = 20$) indicates a high-level of agreement (alpha = 0.82). The table below presents the alpha values if an individual judge is dropped; a resulting alpha value above 0.82 indicates that the dropped judge does not generally agree with the other nineteen judges, whereas a resulting alpha value below 0.82 indicates that the dropped judge contributes to a higher level of agreement among the judges.

For example, the mathematician labeled M2 rated most differently from the other judges, insofar as his removal leads to the sole increased alpha score (alpha = 0.83). On the other hand, the psychologist labeled P4 contributed the most to agreement among the judges, insofar as her removal leads to the minimal resulting alpha score (alpha = 0.79).

Table 1. Adjusted Cronbach's coefficient alpha if a judge is removed.

Reliability if a judge is dropped	
Judge	α
T1	0.82
T2	0.82
T3	0.81
T4	0.82
T5	0.82
M1	0.80
M2	0.83
M3	0.82
M4	0.82
M5	0.80
M6	0.82
P1	0.80
P2	0.81
P3	0.81
P4	0.79
P5	0.82
P6	0.80
P7	0.82
P8	0.81
P9	0.80

Storme et al (2014) remark that alpha scores tend to be higher for a greater number of judges, and cite Kaufman et al (2013) as suggesting that a boot-strapping technique be used when there are groups of judges with different sample sizes. In this study, the psychologists account for nearly twice the number of mathematics teachers; to ameliorate possible sample size problems, the suggested boot-strapping approach was carried out as follows: randomly re-sample sets of judges with replacement, where the total number of judges picked each time is equal to the number of judges in the smallest group, i.e., 5, as corresponds to the number of mathematics teachers. Cronbach's coefficient alpha is then computed for the randomly selected group of five judges. The process is then repeated numerous times and an average alpha value is computed; here, the computations were carried out 100,000 times, which led to a score of alpha = 0.55 with a standard error of 0.17. A score at this level indicates poor intergroup agreement among the three groups of judges.

Intragroup Agreement

To assess the degree of agreement in each of three groups of judges, namely, mathematics teachers, mathematicians, and psychologists, Cronbach's coefficient alpha was computed for each individual group, and then verified by using ICC as well.

Mathematics Teachers

The question addressed here is whether or not there exists intragroup agreement among mathematics teachers with regard to how creativity is conceived of in the context of mathematical problem posing.

A total of five mathematics teachers subjectively rated the creativity of the twenty-five multiplication table problems. Note that the same number of judges is used by Baer (1993) when using CAT to evaluate the creativity of mathematical tasks for eighth-graders (word problem creating, equation creating) and second-graders (word problem creating). Cronbach's coefficient alpha for the group of mathematics teachers ($n = 5$) indicates a remarkably low level of agreement (alpha = 0.16). Table 2 presents the alpha values if an individual teacher is dropped; a resulting alpha value above 0.16 indicates that the dropped teacher does not generally agree with the other four teachers, whereas a resulting alpha value below 0.16 indicates that the dropped teacher contributes to a higher level of agreement among the teachers.

For example, the teacher labeled T2 rated most differently from the other teachers, insofar as her removal leads to the highest resulting alpha score (alpha = 0.28). On the other hand, the teacher labeled T3 contributed the most to agreement among the teachers, insofar as his removal leads to the minimal resulting alpha score (alpha = –0.32). Strictly speaking, alpha values only make sense in the range of 0 to 1, and the negative resulting alpha presented here should be interpreted as meaning that without T3, the agreement among teachers would fall to essentially none at all.

Table 2. Adjusted Cronbach's coefficient alpha if a teacher is removed.

Reliability if a teacher is dropped	
Judge	α
T1	0.20
T2	0.28
T3	-0.32
T4	0.18
T5	0.16

The intraclass correlation method yielded a similar conclusion; in particular, the output was very close to zero (-0.20) which indicates essentially no agreement among the group of mathematics teachers who chose to participate as raters for this study.

Mathematicians

The question addressed here is whether or not there exists intragroup agreement among mathematicians with regard to how creativity is conceived of in the context of mathematical problem posing.

Cronbach's coefficient alpha for the group of mathematicians (n = 6) indicates a poor level of agreement (alpha = 0.48). Table 3 presents the alpha values if an individual mathematician is dropped; a resulting alpha value above 0.48 indicates that the dropped mathematician does not generally agree with the other five mathematicians, whereas a resulting alpha value below 0.48 indicates that the dropped mathematician contributes to a higher level of agreement among the mathematicians.

For example, the mathematician labeled M2 rated most differently from the other mathematicians, insofar as his removal leads to the highest resulting alpha score (alpha = 0.57).

Note that this is also the individual who differed most when all judges were tested together. On the other hand, the mathematician labeled M1 contributed the most to agreement among the teachers, insofar as her removal leads to the minimal resulting alpha score (alpha = 0.13).

Table 3. Adjusted Cronbach's coefficient alpha if a mathematician is removed.

Reliability if a mathematician is dropped	
Judge	α
M1	0.13
M2	0.57
M3	0.49
M4	0.42
M5	0.31
M6	0.55

The intraclass correlation method yielded a very similar conclusion; in particular, the output was the same as the alpha score (0.48) which indicates poor agreement among the group of mathematicians who chose to participate as raters for this study.

Psychologists

The question addressed here is whether or not there exists intragroup agreement among psychologists who work in mathematics education with regard to how creativity is conceived of in the context of mathematical problem posing.

Cronbach's coefficient alpha for the group of psychologists (*n* = 9) indicates a high level of agreement (alpha = 0.78). Table 4 presents the alpha values if an individual psychologist is dropped; a resulting alpha value above 0.78 indicates that the dropped psychologist does not generally agree with the other eight psychologists, whereas a resulting alpha value below 0.78 indicates that the dropped psychologist contributes to a higher level of agreement among the psychologists.

For example, the psychologist labeled P7 rated most differently from the other psychologists, insofar as his removal leads to the highest resulting alpha score (alpha = 0.79). On the other hand, the psychologist labeled P4 contributed the most to agreement among the psychologists, insofar as her removal leads to the minimal resulting alpha score (alpha = 0.71). Note that this is also the individual who contributed the most to agreement when all judges were tested together.

Table 4. Adjusted Cronbach's coefficient alpha if a psychologist is removed.

Reliability if a psychologist is dropped

Judge	α
P1	0.72
P2	0.78
P3	0.75
P4	0.71
P5	0.78
P6	0.75
P7	0.79
P8	0.76
P9	0.74

The intraclass correlation method yielded a very similar conclusion; in particular, the output (0.74) was very close to the alpha score (0.78) which indicates a high level of agreement among the psychologists in mathematics education who chose to participate as raters for this study.

Additional Data

Table 5 contains descriptive statistics for all twenty judges with regard to their respective mean ratings and standard deviations in rating the creativity of the multiplication table problems.

For example, the teacher labeled T1 is an outlier insofar as his mean creativity rating is very close to the maximum of 5. The corresponding standard deviation for T1 is of necessity relatively low, since nearly all of the problems must have been rated with a 5.

The precise creativity ratings for each judge are summarized in Table 6 as the percentage of their responses that were rated a 1, 2, 3, 4, or 5. For example, four of the five teachers (T1, T2, T4, T5) gave out no 1s, and two of the mathematicians (M3, M5) along with three of the psychologists (P6, P7, P9) gave out no 5s. The full five-hundred entry matrix of creativity ratings, with rows corresponding to the twenty-five multiplication table problems and columns corresponding to the judges' respective responses, can be found in Appendix A.

Table 5. Creativity ratings' mean and standard deviation for all judges.

Rating Statistics		
Judge	mean	sd
T1	4.8	0.52
T2	3.6	1.11
T3	3.1	0.93
T4	4.1	0.88
T5	4.4	0.76
M1	3.4	1.47
M2	2.9	1.30
M3	3.5	0.59
M4	3.4	1.00
M5	3.2	1.01
M6	3.0	1.46
P1	3.0	1.40
P2	3.4	0.96
P3	4.0	0.93
P4	3.4	1.35
P5	2.8	1.15
P6	3.0	1.02
P7	3.1	1.32
P8	2.6	0.92
P9	2.7	1.14

It should also be re-iterated that Hennessey et al (2011) suggest presenting the products in a randomized order so as to avoid high levels of agreement appearing as methodological artifacts. In this study, the products were presented in the same order to all 20 judges; however, agreement trends found among-groups and within-groups suggest that this presentation is unlikely to have had a significant effect on the outcomes.

Table 6. Creativity rating by frequency for all judges.

Response Frequency For Each Judge					
	1	2	3	4	5
T1	0.00	0.00	0.04	0.16	0.80
T2	0.00	0.20	0.24	0.28	0.28
T3	0.04	0.20	0.40	0.32	0.04
T4	0.00	0.04	0.20	0.36	0.40
T5	0.00	0.00	0.16	0.28	0.56
M1	0.12	0.24	0.08	0.24	0.32
M2	0.16	0.28	0.20	0.24	0.12
M3	0.00	0.04	0.44	0.52	0.00
M4	0.08	0.04	0.36	0.44	0.08
M5	0.08	0.16	0.20	0.56	0.00
M6	0.28	0.00	0.28	0.28	0.16
P1	0.20	0.16	0.20	0.28	0.16
P2	0.00	0.20	0.32	0.36	0.12
P3	0.00	0.04	0.32	0.28	0.36
P4	0.12	0.16	0.20	0.28	0.24
P5	0.12	0.32	0.28	0.20	0.08
P6	0.12	0.12	0.36	0.40	0.00
P7	0.12	0.24	0.28	0.16	0.20
P8	0.12	0.36	0.36	0.16	0.00
P9	0.16	0.32	0.16	0.36	0.00

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The final chapter is organized into three main sections. The first section is a short summary of the study conducted. The second section answers all four of the research questions set out in the introduction, and follows them with several points of consideration. The chapter concludes with the third section, which provides recommendations for future research based on the results obtained herein.

Summary

Both creativity and problem posing have been important areas of study since at least the mid-twentieth century. Despite the assertion that creativity and problem posing are connected, the precise nature of their relationship remains unclear. Given that each of these two topics is important with regard to mathematics education, this study sought to understand better the creativity involved in mathematical problem posing.

In examining the psychological construct of creativity, one initial approach is to divide into the trichotomy of considering creative people, creative processes, or creative products. A particular advantage in carrying out research on creative products is that they provide concrete objects that can be evaluated in various ways. Often the evaluation of a product with regard to its creativity will be carried out by judges who have relevant domain-expertise; for example, experienced artists might evaluate the creativity of paintings or collages. This approach to evaluating products is part of the Consensual Assessment Technique (CAT), where relevant experts are asked to assess a product’s creativity.

Previous studies using CAT have found a high level of agreement among expert judges from different groups. The goal of this study was to examine mathematical problem posing by beginning with problems, specifically, problems posed using the multiplication table, and to determine whether or not various judges with domain-relevant experience would agree in their subjective assessments of the problems' creativity. The judges recruited for participation were elementary school mathematics teachers, professors or professors emeriti of mathematics, and psychologists who work in mathematics education. Each judge was asked to rate the creativity of twenty-five different multiplication table problems using his or her own subjective definition of creativity; subsequently, within-group agreement for each of the three groups of judges and among-group agreement for all of the judges was tested using a combination of Cronbach's coefficient alpha and the intraclass correlation method, as in earlier studies utilizing CAT.

Conclusions

Research Question 1

Is there intergroup agreement among mathematics teachers, mathematicians, and psychologists as to how creativity is conceived of in the context of mathematical problems posed by prospective mathematics teachers?

This question was investigated using two different methods: first, by inputting all data on creativity ratings for the twenty judges and computing Cronbach's coefficient alpha; second, by using a re-sampling technique on five randomly selected judges (with replacement) whereby Cronbach's coefficient alpha was computed 100,000 times, and an average value was determined.

The first method found a high-level of agreement (alpha = 0.82) and is similar to earlier studies where all judges are evaluated together. However, the tendency of alpha scores to rise with the number of judges brings into question the accuracy of this approach; in particular, the difference in intragroup ratings discussed below combined with the difference in sample sizes suggests a more reasonable portrait is painted by utilizing the second method. To this end, a very low-level of agreement (alpha = 0.55) was found.

The results summarized above cast serious doubts on operationalizing a consensual definition of creativity as a reasonable approach to evaluating products within the given domain among the chosen judges. More precisely, there is insufficient evidence to conclude that there exists intergroup agreement among mathematics teachers, mathematicians, and psychologists with regard to what constitutes a creative mathematical problem posed from the multiplication table.

Research Question 2

Is there intragroup agreement among mathematics teachers with regard to how creativity is conceived of in the context of mathematical problems posed by prospective mathematics teachers?

Cronbach's coefficient alpha for the creativity ratings of five mathematics teachers with experience teaching elementary school mathematics was remarkably low (alpha = 0.16). Though the sample size was smallest for these judges, it is interesting to note that only the removal of T3 would have led to a lower alpha score; removing T5 leaves the alpha score essentially fixed, whereas removing any of the other three teachers would lead to a higher alpha score. Nevertheless, removing the teacher who rated most differently from the others, T2, leads to an adjusted alpha score (alpha = 0.28) that is still far from the lower bound of 0.70 that is used to indicate agreement among judges. Similarly, the ICC score of close to zero indicates a lack of agreement among the mathematics teachers who chose to participate in this study.

The results summarized above cast serious doubts on operationalizing a consensual definition of creativity as a reasonable approach to evaluating products within the given domain among the mathematics teachers. More precisely, there is insufficient evidence to conclude that there exists intragroup agreement among mathematics teachers with regard to what constitutes a creative mathematical problem posed from the multiplication table.

Research Question 3

Is there intragroup agreement among mathematicians with regard to how creativity is conceived of in the context of mathematical problems posed by prospective mathematics teachers?

Cronbach's coefficient alpha for the creativity ratings of six mathematicians was very low (alpha = 0.48). The group of mathematicians contained the single participant, M2, who contributed most to disagreement among all judges as evidenced by the computations found in Table 1. The same phenomenon with regard to M2 is observed among mathematicians, in particular. Nevertheless, even after removing this judge, the adjusted alpha score (alpha = 0.57) remains significantly below the lower threshold of 0.70 used to indicate agreement among judges. Similarly, the same ICC score (0.48) indicates a poor level of agreement among the mathematicians who chose to participate in this study.

The results summarized above cast serious doubts on operationalizing a consensual definition of creativity as a reasonable approach to evaluating products within the given domain among the mathematicians. More precisely, there is insufficient evidence to conclude that there exists intragroup agreement among mathematicians with regard to what constitutes a creative mathematical problem posed from the multiplication table.

Research Question 4

Is there intragroup agreement among psychologists working in mathematics education with regard to how creativity is conceived of in the context of mathematical problems posed by prospective mathematics teachers?

Cronbach's coefficient alpha for the creativity ratings of nine psychologists who worked in mathematics education was high (alpha = 0.78). The group of psychologists contained the single participant, P4, who contributed most to agreement among all judges as evidenced by the computations found in Table 1. The same phenomenon with regard to P4 is observed among psychologists, in particular. Nevertheless, even after removing this judge, the adjusted alpha score (alpha = 0.71) remains above the lower threshold of 0.70 used to indicate agreement among judges. In particular, any eight judge subset of the nine psychologists results in an adjusted alpha score above 0.70; moreover, a single removal among three of the psychologists (P2, P5, P7) leads to an adjusted alpha score that either remains constant or is greater than that of the full group of psychologists. Similarly, the ICC score (0.74) indicates a high level of agreement among the psychologists in mathematics education who chose to participate in this study.

In contrast with the previous three research questions, the results summarized above suggest that operationalizing a consensual definition of creativity is a reasonable approach to evaluating products within the given domain among the psychologists who work in mathematics education. More precisely, there is sufficient evidence to conclude that there exists intragroup agreement among psychologists who work in mathematics education with regard to what constitutes a creative mathematical problem posed from the multiplication table.

Points for Consideration

Given that previous studies have established the reliability and validity of the Consensual Assessment Technique as applied to products in a host of domains, there is a clear and present need to consider more deeply the possible reasons for the outcomes contained herein.

Sample Size

One of the obvious features to emerge in answering the four research questions is that agreement, as measured using Cronbach's coefficient alpha and ICC, was highest when the number of judges was greatest. This section discusses the former, more common measure. Indeed, the cases for mathematics teachers (n = 5), mathematicians (n = 6), and boot-strapping for all three groups by re-sampling five judges repeatedly all indicated low agreement. Conversely, the measurement obtained for the psychologists (n = 9) indicated high agreement, as did testing all the judges (n = 20) together. To this end, at least three points must be addressed.

First, previous everyday creativity studies with similar item and rater counts did not run into analogous difficulties. These studies include art collages as discussed by Amabile (1983) in which the numbers of items and judges, respectively, were 22 and 7 for one study, and 28 and 8 for another. She also cites separate studies on Haikus in which the numbers of item and judges, respectively, were 37 and 4 for one study, and 29 and 4 for another. A more recent study by Storme et al (2014) used 28 drawings and 6 experts. All of these studies report high levels of agreement among the judges; for this study, in which the numbers of items (i.e., multiplication table problems) is 25, the groups with 5 and 6 judges all led to low agreement. Since Cronbach's alpha coefficient is known to increase, generally, with the number of judges, it may make sense to replicate the study with more mathematics teachers and more mathematicians.

Second, in spite of the relatively small sample sizes for two of the three groups, it is worth pointing out that reducing them by a single judge did *not* lower overall agreement in most of the cases. More precisely, for four of the five mathematics teachers, their individual removal leads to an adjusted alpha *greater than or equal to* the one found for all five teachers clumped together. This suggests that the creativity rating patterns are sufficiently different for these four teachers so as to introduce enough noise, individually, to lower the overall agreement among teachers. Analogous remarks can be made about half of the mathematicians: for three of these six judges, removing an individual rater led to an adjusted alpha *strictly greater than* the one found for all six mathematicians clumped together.

Third, for the skeptical reader who persists in the belief that the sample size for mathematics teachers and mathematicians is too low, one approach might be to proceed as in Baer (1993) by combining professors and teachers of mathematics into one group of judges: mathematics instructors (n = 11). Given the validity and reliability pre-established for the Consensual Assessment Technique, it would be reasonable to hypothesize that this larger sample size would lead to better agreement; perhaps even resulting in an alpha value that exceeds that of the 0.78 found for the nine psychologists. However, executing this method on the newly formed group of mathematics instructors still leads to low agreement (alpha = 0.51). Furthermore, Table 7 summarizes the adjusted alpha values when an individual instructor is removed and the test is carried out on the remaining ten instructors. Even when dropping the instructor, M6, who contributes least to overall agreement in this group, the result gives a low level of agreement (alpha = 0.59) that is significantly less than even the least agreeable subset of eight psychologists (alpha = 0.71).

Table 7. Adjusted Cronbach's coefficient alpha if an instructor is removed.

Reliability if an instructor is dropped	
Judge	α
T1	0.50
T2	0.54
T3	0.43
T4	0.49
T5	0.52
M1	0.38
M2	0.48
M3	0.51
M4	0.48
M5	0.35
M6	0.59

Thus, while replicating this study with more mathematics teachers, more mathematicians, or more of both may be of interest, it is unlikely that the low agreement found here can be attributed simply to issues around sample size.

Thought Processes

An additional area of consideration is how judges in different groups think about the multiplication table problems that they were asked to rate. In this respect, one might wonder about how the judges view problems posed at this level of complexity, in particular, as well as how they think about mathematics, more generally.

The elementary school mathematics teachers who participated in this study all have experience with either teaching the multiplication table or with teaching students who covered the multiplication table relatively recently. A reasonable hypothesis is that, in keeping with the general way in which the multiplication table is taught, most teachers will not have explored the sort of nonstandard problems presented as a part of this study for this particular mathematical scenario. One place to look for evidence of such a phenomenon is in the average creativity scores assigned to the problems. Of the twenty judges in total, precisely four of them had mean creativity ratings that were greater than or equal to 4.0; three of these four judges are found among the mathematics teachers. This information is summarized in Table 5; more precisely, the judges and their respective mean ratings are: T1, 4.8; T4, 4.1; T5, 4.4; and P3, 4.0. Rephrased, 3 out of 5 teachers had mean creativity ratings *strictly greater than* 4.0, whereas all 15 of the mathematicians and psychologists had mean ratings *less than or equal to* 4.0. While the relative rating trends indicate a lack of agreement, the descriptive statistics indicate an important way in which the teachers *did* agree: namely, they found almost all of the problems to be very creative. When these stakeholders find problems to be creative almost without exception, it suggests that they do not consider present materials for the multiplication table to be sufficiently creative.

The mathematicians who participated in this study are unlikely to have taught the multiplication table at the tertiary level, and are unlikely to have had students who recently covered it in a course. Gruber and Wallace (1999) wonder in their discussion of creative individuals: “Do mathematicians think in equations?” (p. 103). The question of how mathematicians think is not so easily answered. Hadamard (1945) remarks that nearly all of the mathematicians who he knows “avoid not only the use of mental words but also, just as I do, the mental use of algebraic or any other precise signs; also as in my case, they use vague images” (p. 84). He contrasts this method of thinking with a few notable examples: George D. Birkhoff, “who is accustomed to visualize algebraic symbols and to work with them mentally,” Norbert Wiener, who “happens to think either with or without words,” and Jesse Douglas, whose “research thought is in connection with words, but only with their rhythm, a kind of Morse language where only the numbers of syllables of some words appear” (p. 84). Hadamard goes on to discuss the curious case of George Polya, already mentioned as the founder of mathematical problem solving as an area of research, who is the rare exception as a mathematician in that his thoughts rely on words and letters. Hadamard quotes Polya describing his own thinking: “I believe… that the decisive idea which brings the solution of a problem is rather often connected with a well-turned word or sentence… The right word, the subtly appropriate word, helps us to recall the mathematical idea…” (p. 84). Hadamard himself continues: “Moreover, [Polya] finds that a proper notation – that is, a properly chosen letter to denote a mathematical quantity – can give him similar help; and some kind of puns, whether of good or poor quality, may be useful for that purpose” (p. 85). He also discusses physical intuition and mental pictures of a kinetic sort as being relevant to mathematicians’ ways of thinking.

How would Polya have regarded a question about monosyllabic number names in the multiplication table? How would the mathematicians thinking with visual images respond to questions based upon a different mathematical scenario, such as one rooted in a geometric foundation? Though the remarks of Hadamard correspond to the creativity of mathematicians themselves, one might also wonder how such modalities of thought manifest in mathematicians’ evaluation of the creativity of other individuals or items. As the domain of mathematics becomes increasingly specialized, it may be that mathematicians who focus in a particular area or have a particular way of thinking about mathematics are inclined to interpret the creativity of mathematical problems in very different manners.

Lastly, psychologists who conduct research in mathematics education are the sole group among the three who have almost certainly been presented with, and perhaps even distributed, surveys in which psychological constructs are rated or evaluated. Their familiarity with Likert scales and studies conducted in the domain of mathematics education contrasts with the experiences of many elementary school mathematics teachers and professors of mathematics. It is unclear, but perhaps worthy of investigation, as to whether this shared background contributed to the high-level of agreement found in their creativity ratings. In addition, further commentary about psychologists' domain-expertise is warranted; such remarks are deferred until later on in the discussion section.

Problem Complexity

In discussing the validity and reliability of the Consensual Assessment Technique, Baer and McKool (1993) note that earlier studies about collages found ratings of complexity correlated highly with ratings of creativity. As the authors write, complexity is an aspect "of a collage that *should* be related to the creativity of that collage" (p. 5). Insofar as mathematical problem posing by prospective teachers is concerned, does the analogous normative statement hold? In particular, should the complexity of a posed problem correlate highly with its creativity?

From a theoretical vantage point, some domains use language that is complex to the extent that it is inaccessible to most. As Csikszentmihalyi (1999) writes: "The autonomy of a field is to a certain extent a function of the codification of the domain it serves. When the domain is arcane and highly codified, like Assyriology or molecular biology, then the decision as to which new meme is worth accepting will be made by a relatively small field that is committed to following the traditions and rules of the domain" (p. 326). If molecular biology is to be admitted as "arcane and highly codified," then certainly mathematics ought to be as well. With regard to the problems generated for this study, the mathematical language utilized was of necessity relatively simple, for the questions were intended to be posed to elementary school students who had covered the multiplication table. Nevertheless, there is a *conceptual complexity* to the problems, particularly with regard to how they are solved.

As a concrete example, consider the first problem posed: "Estimate how many distinct numbers are in the 100 boxes of a 10x10 multiplication table." At the elementary school level, observing the diagonal symmetry of a multiplication table and guessing around 50 ought to be considered a reasonable answer by students. Some may note that many of the numbers appear multiple times. For example, 24 appears four times: 3*8, 4*6, 6*4, and 8*3. Bearing this in mind, a more precise estimation of slightly less than 50 might be provided by advanced students; indeed, the precise number of distinct entries is 42. A natural generalization, however, is to ask the same question about an *n* by *n* table. Both Problem 9 ("Using a multiplication table, explain what it means for a number to be prime") and Problem 17 ("Which numbers are missing from the 10x10 multiplication table and why?") suggest a relation between prime numbers and the multiplication table. (In particular, a number is prime if and only if it appears precisely twice in an indefinitely extended multiplication table.) From this vantage point, it comes as less of a surprise that the generalized form of Problem 1 is intractable, even at the advanced undergraduate level. This generalization is related to Erdos' multiplication table problem, due to the number theorist Paul Erdos (1965), and was settled only in the twenty-first century by Ford (2008). What, then, can be said about the *complexity* of the first multiplication table problem distributed, and how might it relate to the problem's perceived *creativity*?

The aforementioned example is not an anomaly. In all likelihood, the symbols introduced within the first year of an undergraduate mathematics major's coursework are codified enough to be indecipherable at a glance for all but the mathematicians. Manifestations of this codification range from a triple integral in polar coordinates as discussed in a course on multivariable Calculus, to Euler's totient function as introduced in a first course on Number Theory. Yet questions about these mathematical objects can be asked for which the answer is readily provided by even a strong secondary school student. Meanwhile, consider questions such as the following: "Can every even number from 4 on be written as a prime plus another prime?" or

“Starting with any whole number: Multiply by 3 and add 1 if it is odd, and divide by 2 if it is even. Repeating this process indefinitely, will the number 1 necessarily appear?” These problems, though understandable as they are *posed*, are as of yet *unsolved*. This is not without attempts made by the field of mathematicians: an answer of yes to the former is known as Goldbach’s Conjecture, and an answer of yes to the latter is known as the Collatz Conjecture; both are well-known among mathematicians as intensely difficult, open problems. The phenomenon of a question that is simple to *pose* but difficult to *solve* presents itself yet again.

The lesson here may be that, just as problem posing and problem solving were presented early on as being related, in fact the two are inextricably linked. Under such an interpretation, the task of rating posed problems would necessitate the consideration of how they could be solved. In this respect, the complexity of the problem is in some part a function of the mathematics used to resolve it; moreover, it seems reasonable to assume that elementary school mathematics teachers, mathematicians, and psychologists would be inclined to solve (or attempt to solve) the multiplication table problems in different ways, and, possibly, to consider various generalizations or specifications. To this end, a possible consequence is that dissimilar approaches might lead to dissimilar ratings among the judges.

Defining Domains

A subtle difficulty in recruiting judges with domain-expertise is first to specify the domain itself. In her early work on the Consensual Assessment Technique, Amabile (1983) writes: “These data, of course, raise the question of who is to be considered an ‘expert’ for the purposes of the present methodology and, thus, who is to be considered an ‘appropriate’ judge. It appears that the only requirement is a familiarity with the domain of endeavor in which the product was made” (p. 61). In a study on the problem posing of prospective mathematics teachers, what exactly is the domain? Comparing it to earlier studies on collages would lead one to believe that the domain ought to be mathematics. Csikszentmihalyi (1999) would likely agree with this specification; the domain examples he provides include: religion, philosophy, mathematics, music, and technology. These are all large domains, which could easily be conceived of as containing smaller domains within them; for the aforementioned examples, respective subdomains could be: Jewish studies, metaphysics, algebra, piano-playing, and software development.

However, Csikszentmihalyi also includes seemingly smaller domains, such as Assyriology (as opposed to linguistics or history) and woodworking (as opposed to craftsmanship). Though he gives no precise way of specifying what does or does not constitute a domain, he does remark briefly about Piaget's (1965) work on a Swiss marble game:

> It is usually the case that, with time, a domain develops its own memes and systems of notation. Natural languages and mathematics underlie most domains. In addition there are formal notation systems for music, dance, and logic, as well as less formal ones for instructing assessing performance in a great variety of different domains. For instance, Jean Piaget gave a detailed description of how rules are transmitted in a very informal domain, that of the game of marbles played by Swiss children. This domain has endured over several generations of children, and it consists of specific names for marbles of different sizes, colors, and composition (p. 319).

This quotation is perplexing insofar as mathematics has been mentioned as both a domain and as underlying the majority of domains; furthermore, the Swiss marble game is referred to first as a "very informal domain" and then just a "domain," and as having its own system of notation. In the absence of a strict definition for domains, one can only infer that it must be broad enough to allow both *mathematics* and *a particular marble game played by Swiss children* to be admitted. Moreover, there exists an implicit notion of formality of domains; a topic into which Csikszentmihalyi does not delve further.

For the study at hand, the relevant judges were chosen without explicitly specifying the domain. Whether the domain is mathematics, education, or mathematics education, it seems as if Amabile's criterion for choosing appropriate judges is satisfied. But what if the true domain of study here is less broad? Perhaps the domain is elementary school mathematics, or elementary school mathematics teaching, or even elementary school mathematics teaching of multiplication table problems. Certainly these last suggestions are more accurate descriptions of the intended area of study, but does any one of them constitute a domain? If so, then perhaps mathematicians are not an ideal choice as relevant judges. The elementary school teachers would still be relevant, but perhaps a higher bar for teaching expertise should be set, such as picking some sort of *master teachers* or teachers with more years of experience. It may also be the case that the teachers lack certain characteristics associated with creative potential, as found in the Silver and Leung (1997) study with the TTCT-V, which could influence their own creativity ratings. Meanwhile, the one group that is most qualified with respect to any of the domains is almost certainly the psychologists: all are established academics with experience conducting research in mathematics education, and they are more likely to have a broad knowledge of the diverse sorts of problem that can be used to teach the multiplication table. From this perspective, it is to be expected from previous research on the Consensual Assessment Technique in other domains that they would exhibit a high-level of agreement; indeed, among all groups, Cronbach's coefficient alpha and the intraclass correlation method indicated high-agreement only for the psychologists.

Recommendations

Designing programs to foster creativity among prospective mathematics teachers requires that the groups involved in such an endeavor have some shared understanding of what constitutes 'creative' work. Previous studies using the Consensual Assessment Technique demonstrated that, assuming a familiarity with the appropriate domain, subjective agreement about the creativity of products is high enough to let experts' personal views on creativity provide an operational definition. The study here suggests that for the case in mathematics education, and, specifically, with regard to mathematical problems posed using the multiplication table, within-group and among-group agreement for the participants recruited was sufficient to allow the operationalization of subjective views of creativity *only* in the case of psychologists who work in mathematics education.

Whereas previous documents from the National Council for Teachers of Mathematics (NCTM, 1989; NCTM 1991) placed an emphasis on problem posing, the recent Common Core State Standards for Mathematics (Common Core State Standards Initiative, 2010) include a newfound emphasis on mathematical modeling. Modeling is included as a Standard for Mathematical Practice for all grades K–12. It is discussed in greater depth as a strand for secondary school students:

> Real-world situations are not organized and labeled for analysis; formulating tractable models, representing such models, and analyzing them is appropriately a creative process. Like every such process, this depends on acquired expertise as well as creativity (p. 72).

Moreover, an ill-structured problem space, in the sense of creativity theorist Stokes (2005), is encountered by students and teachers who begin with a real-world problem and attempt to model it mathematically. Also in keeping with Stokes' work is the language around constraints, as the Standards continue:

> How precise an answer do we want or need? What aspects of the situation do we most need to understand, control, or optimize? What resources of time and tools do we have? The range of models that we can create and analyze is also constrained by the limitations of our mathematical, statistical, and technical skills, and our ability to recognize significant variables and relationships among them (p. 72).

Thus, the scenario faced in mathematical modeling exemplifies a situation in which one must be creative, and in which one poses certain problems in order to move through the modeling cycle. In spite of the present dearth of curricular materials for teaching mathematical modeling, as well as teacher misconceptions about mathematical models and modeling (Gould, 2013), it may be difficult for mathematics teachers, mathematicians, and psychologists to collaborate on developing curricula or implementing professional development programs to address these current concerns in mathematics education.

More generally, the lack of intergroup agreement found in this study indicates a need for programs to allow the three individual groups to arrive at a better-shared conception of what constitutes creativity. The precise manner in which this need can and should be addressed remains enigmatic; nevertheless, the suggestion here is that it is an important area of future research. Providing opportunities for communication between the different groups to discuss their conceptions of creativity, whether with regard to mathematical problem posing or mathematics more generally, may be a reasonable first step. Moreover, this study reaffirms the importance of domain-expertise in *mathematics education* among influential stakeholders.

The scope of this study was limited to twenty total judges, the topic of mathematical problem posing, and the starting scenario of the multiplication table. Any one of these three features could be varied in order to carry out future research. Recruiting more judges, possibly with different criteria for selection, and having them rate more problems is one way in which to proceed. Alternatively, further studies on other areas of school mathematics would be of interest. For example, given the relations between problem posing and problem solving, one could ask judges to rate the problems' creativity before and after solving them; this could also be compared to the correctness of the solutions provided. As for the multiplication table, similar studies could be carried out using different starting scenarios. Studies specifically on real-world mathematical modeling and its relation to creativity could be of interest and would align closely with the current set of Standards; alternatively, one could compare creative conceptions with regard to problems posed from specific areas of mathematics, such as algebra and geometry at the primary or secondary level, or analysis and discrete mathematics at the tertiary level.

Finally, the study suggests a further investigation of both judges' thought processes as they rate the creativity of mathematical problems, and further theoretical development around the concept of a domain. With regard to the former consideration, one way to proceed would be to follow creativity ratings by obtaining qualitative data from the judges, for example, using the clinical interview technique (Ginsburg, 1997). Alternatively, one might allow the judges from the same or different groups to rate the creativity of mathematical objects, discuss their ratings with one another, and then rate the creativity of the same set (or a new set) of mathematical objects post-discussion. It is not clear how the pre-discussion and post-discussion ratings of creativity would vary.

With regard to the definition of domains, an initial assumption in carrying out research using the Consensual Assessment Technique is that the judges are chosen for their familiarity with the domain. In this respect, the study here is further along the little-c to Big-C continuum than earlier work on collages, Haikus, or equation writing; correspondingly, the results indicate a higher level of domain-expertise was necessary in order to operationalize a consensus definition of creativity. Judges are chosen as representatives of a specific field, and the field itself can be viewed as a function of its corresponding domain. The domain in this study can be conceived of as mathematics or mathematics education without endangering the judge choices; however, a more precise theory of domains might allow future studies to be sharpened in a meaningful way.

The gatekeepers in mathematics and mathematics education include sources other than the three groups discussed here; for example, both school administrators and parents have an important effect on students' mathematical learning outcomes (Useem, 1992). By choosing an appropriate domain, whether a broad one that underlies other domains, such as mathematics, or a more narrowly defined informal domain, such as mathematical problem posing using the multiplication table, researchers could better equip themselves to select judges with the expertise necessary to reach a consensus in their ratings of the creativity of the products under consideration.

REFERENCES

Abramovich, S. (2007). Uncovering hidden mathematics of the multiplication table using spreadsheets. *Spreadsheets in Education (eJSiE)*, *2*(2), 1.

Almeida, L. S., Prieto, L. P., Ferrando, M., Oliveira, E., & Ferrándiz, C. (2008). Torrance Test of Creative Thinking: The question of its construct validity. *Thinking Skills and Creativity*, *3*(1), 53-58.

Amabile, T. M. (1983). A consensual technique for creativity assessment. *The social psychology of creativity*, 37-63. Springer US.

Amabile, T. M. (1996). *Creativity in context: Update to the social psychology of creativity.* Boulder, CO: Westview.

Amabile, T. M. (2012). *Componential theory of creativity*. Harvard Business School.

Baer, J. (1993). Divergent thinking and creativity: A task-specific approach. *Lawrence Erlbaum Associates, Hillsdale, NJ*.

Baer, J. (1998). The case for domain specificity of creativity. *Creativity Research Journal, 11*(2), 173-177.

Baer, J., & McKool, S. (2009). Assessing creativity using the consensual assessment. *Handbook of assessment technologies, methods, and applications in higher education*. Hershey, Pennsylvania: IGI Global.

Bateson, M. C. (2001). *Composing a life*. Grove Press.

Bateson, M. C. (2004). Composing a Life Story. *Willing to learn: Passages of personal discovery* (pp. 66-76). Hanover, NH: Steerford Press.

Balka, D. S. (1974). Creative Ability in Mathematics. *Arithmetic Teacher, 74*.

Barron, F. X., Montuori, A., & Barron, A. (Eds.). (1997). *Creators on creating: Awakening and* *cultivating the imaginative mind*. Putnam.

Beghetto, R. A., & Kaufman, J. C. (2007). Toward a broader conception of creativity: A case for "mini-c" creativity. *Psychology of Aesthetics, Creativity, and the Arts, 1*(2), 73.

Brower, R. (2003). Constructive repetition, time, and the evolving systems approach. *Creativity Research Journal*, *15*(1), 61-72.

Brown, S. I., & Walter, M. I. (1990). The art of problem posing.

Brown, S. I., & Walter, M. I. (Eds.). (2014). *Problem posing: Reflections and applications*. Psychology Press.

Cai, J. (1998). An investigation of US and Chinese students' mathematical problem posing and problem solving. *Mathematics Education Research Journal*, *10*(1), 37–50.

Cai, J., & Hwang, S. (2002). Generalized and generative thinking in US and Chinese students' mathematical problem solving and problem posing. *The Journal of Mathematical Behavior*, *21*(4), 401-421.

Callahan, C. M. (1991). The assessment of creativity. *Handbook of gifted education*, 219-235.

Cantor, G. (1867). *De aequationibus secundi gradus indeterminatis*. C. Schultze.

Chase, W.I. (1985). Review of the Torrance Tests of Creative Thinking. *Ninth mental measurements yearbook*, *2*, 1630-1634. Lincoln, NE: Buros Institute of Mental Measurement.

Cohen, P. (2002). The discovery of forcing. *Rocky Mountain Journal of Mathematics*, 32(4).

Cohen, P. J. (2008). *Set theory and the continuum hypothesis*. Dover Publications.

Colangelo, N., & Davis, G. A. (1997). *Handbook of gifted education*. Allyn & Bacon.

Common Core State Standards Initiative. (2010). *Common Core State Standards for*

Mathematics. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers.

Cropley, A. (2010). Creativity in the classroom: The dark side. *The dark side of creativity*, 297-316.

Csikszentmihalyi, M. (1999). Implications of a Systems Perspective for the Study of Creativity. *Handbook of creativity*, 313-335.

Davis, G. A. (1997). Identifying creative students and measuring creativity. *Handbook of gifted education*, *2*, 253-281.

Dawkin, R. (1976). The selfish gene. *Oxford University Press*, *1*, 976.

Dickman, B. (2013). Mathematical Creativity, Cohen Forcing, and Evolving Systems: Elements for a Case Study on Paul Cohen. *Journal of Mathematics Education at Teachers College*, *4*(2).

Duncker, K., & Lees, L. S. (1945). On problem-solving. *Psychological monographs*, *58*(5).

Erdös, P. (1965). Some remarks on number theory. *Israel Journal of Mathematics*, 3(1), 6-12.

Ford, K. (2008). The distribution of integers with a divisor in a given interval. *Annals of Mathematics*, 367-433.

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

Getzels, J. W. (1975). Problem-finding and the inventiveness of solutions. *The Journal of Creative Behavior*, *9*(1), 12-18.

Ginsburg, H. (1997). *Entering the child's mind: The clinical interview in psychological research and practice*. Cambridge University Press.

Gould, H. (2013). *Teachers' Conceptions of Mathematical Modeling*. Doctoral dissertation, Columbia University.

Gruber, H. E. (1981). On the Relation Between Aha Experiences and the Construction of Ideas in Innovation and Continuity in Science. *History of Science Cambridge*, *19*(1), 41-59.

Gruber, H. E. (1988). The evolving systems approach to creative work. *Creativity Research Journal*, *1*(1), 27-51.

Gruber, H. E., & Barrett, P. H. (1974). *Darwin on man: A psychological study of scientific creativity*. EP Dutton.

Gruber, H. E., & Davis, S. N. (1988). 10 Inching our way up Mount Olympus: the evolving-systems approach to creative thinking. *The nature of creativity: Contemporary psychological perspectives*, *243*.

Gruber, H. E., & Wallace, D. B. (1999). The case study method and evolving systems approach for understanding unique creative people at work. *Handbook of creativity*, *93*, 115.

Guilford, J. P. (1950). Creativity. *American Psychologist*, *5*(9), 444-454.

Guilford, J. P. (1956). The structure of intellect. *Psychological bulletin*, *53*(4), 267.

Guilford, J. P. (1959). Traits of creativity. *Creativity and its cultivation*, 142-161.

Guilford, J. P., & Lacey, J. I. (1947). Printed classification tests.

Hadamard, J. (1945). The Psychology of Mathematical Invention.

Hennessey, B. A. (2003). The social psychology of creativity. *Scandinavian Journal of Educational Research*, *47*(3), 253-271.

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

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

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

Kaufman, J. C., & Sternberg, R. J. (Eds.). (2006). *The international handbook of creativity*. Cambridge University Press.

Kilpatrick, J. (1987). Problem formulating: Where do good problems come from?. *Cognitive science and mathematics education*, 123-147.

Kim, K. H. (2006). Can we trust creativity tests? A review of the Torrance Tests of Creative Thinking (TTCT). *Creativity research journal*, *18*(1), 3-14.

Kunen, K. (1980). *Set theory* (Vol. 313). Amsterdam: North-Holland.

Leung, S. S. (1993). The relation of mathematical knowledge and creative thinking to the mathematical problem posing of prospective elementary school teachers on tasks differing in numerical information content. Unpublished doctoral dissertation, University of Pittsburgh.

Leung, S. S., & Silver, E. A. (1997). The role of task format, mathematics knowledge, and creative thinking on the arithmetic problem posing of prospective elementary school teachers. *Mathematics Education Research Journal*, *9*(1), 5-24.

Lobman, C., & Lundquist, M. (2007). *Unscripted learning: Using improv activities across the K-8 curriculum*. Teachers College Press.

Lubart, T. I. (1990). Creativity and Cross-Cultural Variation. *International Journal of Psychology, 25*(1), 39-59.

Lubart, T. (2010). *Cross-cultural perspectives on creativity*. The Cambridge handbook of creativity, 265-278.

Maslow, A. H. (1963). *The creative attitude*. Psychosynthesis Research Foundation.

Mason, J. H. (2003). *The value of creativity: The origins and emergence of a modern belief*. Ashgate Publishing, Ltd..

May, R. (1974). The courage to create. New York: W. W. Norton & Company.

Moore, G. H. (1988). The origins of forcing, Logic Colloquium'86 (Frank R. Drake and John K. Truss, editors). *Studies in Logic and the Foundations of Mathematics*, North-Holland, Amsterdam, 143-173.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA.

Piaget, J. (1965). *The moral judgment of the child*. New York: Free Press.

Plucker, J. A. (1998). Beware of simple conclusions: The case for content generality of creativity. *Creativity Research Journal, 11*(2), 179-182.

Poincaré, H. (1982). The Foundations of Science. Science and Hypothesis, The Value of Science. *Science and Method, Univ. Press of America*.

Pólya, G. (1945). How to solve it. Princeton. *New Jersey: Princeton University*.

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

Runco, M. A. (1991). *Divergent thinking*. Ablex Publishing.

Runco, M. A. (1993). Divergent thinking, creativity, and giftedness. *Gifted Child Quarterly, 37*(1), 16-22.

Runco, M. A. (Ed.). (1994). *Problem finding, problem solving, and creativity*. Greenwood

Publishing Group.

Runco, M. A. (2008). *Commentary: Divergent thinking is not synonymous with creativity*.

Runco, M. A. (2010). *Divergent thinking, creativity, and ideation*. The Cambridge handbook of creativity, 413-446.

Runco, M. A., Millar, G., Acar, S., & Cramond, B. (2010). Torrance tests of creative thinking as predictors of personal and public achievement: A fifty-year follow-up. *Creativity Research Journal*, *22*(4), 361-368.

Sawyer, R. K., & DeZutter, S. (2009). Distributed creativity: How collective creations emerge from collaboration. *Psychology of Aesthetics, Creativity, and the Arts*, *3*(2), 81.

Sawyer, R. K. (2011). *Structure and Improvisation in Creative Teaching*. Cambridge University Press. 32 Avenue of the Americas, New York, NY 10013.

Schoenfeld, A. H. (1979). Explicit heuristic training as a variable in problem-solving performance. *Journal for Research in Mathematics Education*, 173-187.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. New York: Academic press.

Schoenfeld, A. H. (1987). What's All the Fuss About Metacognition. *Cognitive science and mathematics education*, *189*.

Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*(5), 484.

Shuk-kwan, S. L. (1997). On the role of creative thinking in problem posing. *ZDM*, *29*(3), 81-85.

Silver, E. A. (1979). Student perceptions of relatedness among mathematical verbal problems. *Journal for research in mathematics education*, 195-210.

Silver, E. A. (1994). On Mathematical Problem Posing. *For the learning of mathematics*, *14*(1), 19-28.

Silver, E. A., & Smith, J. P. (1980). Think of a related problem. *Problem Solving in School Mathematics, NCTM, Reston, Virginia*, 146-156.

Simon, H. A. (1979). *Models of thought (Vol. I)*. New Haven: Yale University Press

Sriraman, B. (2004). The Characteristics of Mathematical Creativity. *Mathematics Educator*, *14*(1), 19-34.

Starko, A. J. (2004). Chapter 5: Teaching creative thinking skills and habits; Chapter 7: Motivation, creativity and classroom organization. *Creativity in the classroom: Schools of curious delight*, 177-258; 357-418). Mahway, NJ: Lawrence Erlbaum.

Sternberg, R. J. (Ed.). (1999). *Handbook of creativity*. Cambridge University Press.

Stokes, P. D. (2005). *Creativity from constraints: The psychology of breakthrough*. Springer Publishing Company.

Stokes, P. D. (2010). Using constraints to develop creativity in the classroom. *Nurturing creativity in the classroom*, 88-112.

Storme, M., Myszkowski, N., Çelik, P., & Lubart, T. (2014). Learning to judge creativity: The underlying mechanisms in creativity training for non-expert judges. *Learning and Individual Differences, 32*, 19-25.

Stoyanova, E., & Ellerton, N. F. (1996). A framework for research into students' problem posing in school mathematics. *Technology in mathematics education. Mel bourne: Mathematics Education Research Group of Australia*.

Torrance, E. P. (1968). *Torrance tests of creative thinking*. Personnel Press, Incorporated.

Torrance, E. P. (1979). *The search for satori and creativity*. Buffalo, NY: Creative Education Foundation.

Torrance, E. P. (1988). The nature of creativity as manifest in its testing. *The nature of creativity*, 43-75.

Torrance, E. P. (1998). *Torrance tests of creative thinking: Norms-technical manual: Figural (streamlined) forms A & B*. Scholastic Testing Service.

Torrance, E. P. (1990). *Torrance Tests of Creative Thinking: Figural (streamlined) Forms A & B. Norms-technical Manual*. Scholastic Testing Service, Incorporated.

Torrance, E. P., & Safter, H. T. (1990). *The incubation model of teaching: Getting beyond the aha!*. Bearly Limited.

Treffinger, D. J. (1985). Review of the Torrance Tests of Creative Thinking. *Ninth mental measurements yearbook*, *2*, 1632-1634. Lincoln, NE: Buros Institute of Mental Measurement.

Treffinger, D. J., & Selby, E. C. (1993). Giftedness, creativity and learning style: exploring the connections. *Teaching and counseling gifted and talented adolescents: An international learning style perspective*, 87-102.

Trivett, J. (1980). The Multiplication Table: To Be Memorized or Mastered?. *For the Learning of Mathematics*, 21-25.

Useem, E. L. (1992). Getting on the fast track in mathematics: School organizational influences on math track assignment. *American journal of Education*, 325-353.

Van Harpen, X. Y., & Sriraman, B. (2013). Creativity and mathematical problem posing: an analysis of high school students' mathematical problem posing in China and the USA. *Educational Studies in Mathematics*, *82*(2), 201-221.

Wallas, G. (1926). *The art of thought*. New York: Harcourt, Brace and Company.

Weisberg, R. W. (1988). 6 Problem solving and creativity. *The nature of creativity:*

Contemporary psychological perspectives, 148.

Weisberg, R. W. (2004). On structure in the creative process: A quantitative case-study of the creation of Picasso's Guernica. *Empirical Studies of the Arts*, *22*(1), 23-54.

Weisberg, R. W. (2006). *Creativity: Understanding innovation in problem solving, science, invention, and the arts*. John Wiley & Sons.

Weisberg, R. W. (2011). Frank Lloyd Wright's Fallingwater: A case study in inside-the-box creativity. *Creativity Research Journal*, *23*(4), 296-312.

Wertheimer, M. (1959). *Productive thinking: Enlarged Edition*. New York: Harper.

APPENDIX A

CREATIVITY RATINGS

	T1	T2	T3	T4	T5	M1	M2	M3	M4	M5	M6	P1	P2	P3	P4	P5	P6	P7	P8	P9
Q1	5	3	4	2	5	3	2	4	4	3	4	5	3	3	3	5	4	4	2	4
Q2	5	3	2	3	5	5	2	4	3	4	3	3	5	3	2	3	1	4	3	3
Q3	4	5	3	4	5	5	2	4	3	4	4	4	4	5	5	4	4	5	4	3
Q4	5	5	4	4	4	2	2	3	4	4	5	5	3	5	5	4	4	2	4	4
Q5	4	3	2	5	4	2	1	4	3	1	3	1	3	4	1	3	1	3	3	1
Q6	5	4	2	3	5	1	2	3	1	1	1	1	2	3	1	3	3	3	2	1
Q7	3	2	3	3	5	2	3	3	4	3	3	1	3	2	3	1	1	2	3	3
Q8	5	2	4	5	4	1	4	3	1	4	3	4	3	3	4	2	4	2	2	2
Q9	5	5	2	5	3	4	1	3	3	4	4	2	4	3	4	4	3	1	2	3
Q10	5	4	3	5	5	5	2	4	4	4	5	4	4	4	4	3	3	5	3	4
Q11	5	4	3	5	5	5	3	4	4	4	5	4	4	5	4	3	4	5	3	4
Q12	5	2	4	5	5	4	2	3	3	4	4	3	2	3	3	2	3	3	2	4
Q13	5	4	1	3	3	5	4	4	4	2	4	2	2	3	1	3	3	1	1	1
Q14	5	5	2	4	5	1	1	3	4	2	1	4	4	4	2	1	2	2	1	2
Q15	5	5	4	3	4	2	1	4	3	4	5	4	3	5	4	2	2	5	4	2
Q16	5	4	4	4	5	4	4	4	2	4	4	5	3	5	5	4	4	4	2	4
Q17	5	2	3	4	3	5	5	3	4	4	1	5	4	4	5	3	4	2	2	2
Q18	5	2	3	4	3	2	3	3	3	3	3	1	2	5	3	2	3	3	3	2
Q19	5	3	3	4	4	4	3	4	3	2	3	2	4	5	4	2	3	1	3	2
Q20	4	5	3	5	4	2	4	4	4	2	1	3	2	4	2	5	3	3	1	1
Q21	5	5	3	5	5	5	3	3	4	4	4	4	4	5	5	4	4	3	4	4
Q22	4	3	4	4	4	4	5	3	3	3	1	1	4	3	2	1	2	5	3	4
Q23	5	3	5	5	5	5	5	2	5	4	1	3	5	5	4	2	4	2	3	4
Q24	5	4	4	4	5	3	4	4	4	3	3	3	3	4	5	2	4	3	2	2
Q25	5	4	3	5	5	4	4	4	5	4	1	2	5	4	3	2	3	4	2	2

APPENDIX B

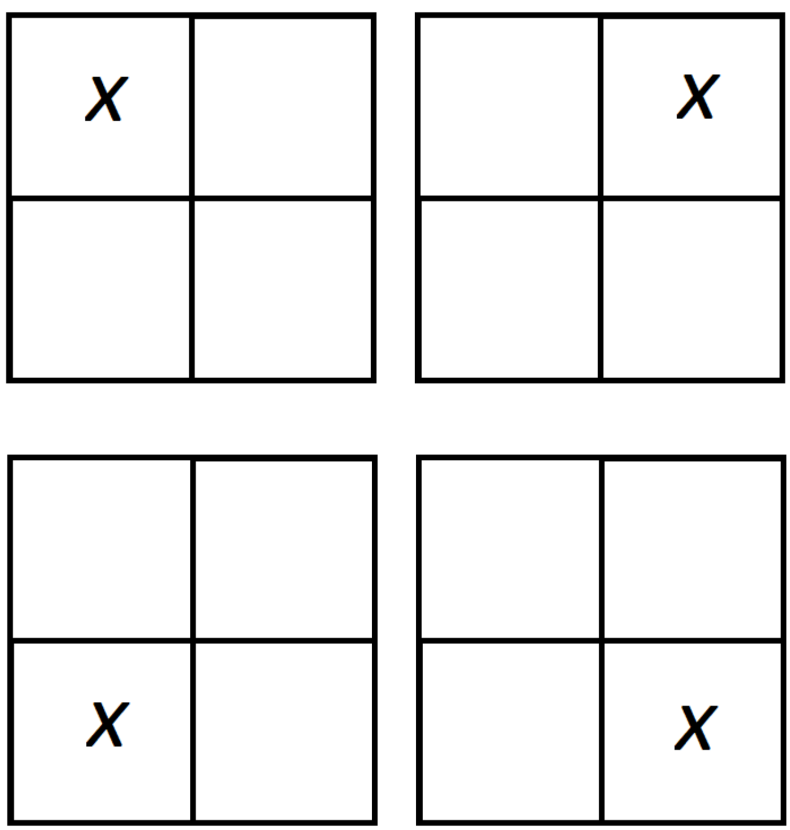
MULTIPLICATION TABLE PROBLEMS

1. Estimate how many distinct numbers are in the 100 boxes of a 10x10 multiplication table.

2. How many of the entries in the 10x10 multiplication table are odd?

3. Starting at the 1 in a multiplication table, you can take steps to adjacent boxes: right or left, up or down, but not diagonally. What is the biggest number you can arrive at in 10 steps?

4. Consider the 2x2 sub-grid cutouts from a multiplication table. For which *x* can you fill in the remaining three numbers with complete surety?



5. Find the only 49 in the multiplication table. Add to it the six numbers above in its column, and the six numbers to the left in its row. What is the total?

6. Add up the numbers from all 100 boxes in a 10x10 multiplication table. What is the total?

7. Where do the multiples of 3 appear in the 10x10 multiplication table?

8. Which operations (Addition? Subtraction? Multiplication? Division? Other?) are related to the 10x10 times table?

9. Using a multiplication table, explain what it means for a number to be “prime.”

10. Look at the numbers on the south-east diagonal starting from either 2 entry in the 10x10 times table: 2, 6, 12, 20, 30, 42, 56, 72, 90. What pattern or patterns do you see?

11. Starting at the 3 in row one of the multiplication table, jump like a knight (from Chess) to move one square down and two to the right. Doing this repeatedly, you get the following numbers: 3, 10, 21, 36. What pattern or patterns do you see?

12. What kinds of symmetry can you find in the 10x10 multiplication table?

13. How many two digit numbers are in the 10x10 multiplication table?

14. Suppose you have a 10x10 table of blank entries and you want to fill it out as a multiplication table. In what order would you write in the numbers?

15. In the 10x10 multiplication table, the number directly to the right of x is $x+8$. What is the next number over to the right?

16. How would you extend the 10x10 multiplication table to cover the negative numbers?

17. Which numbers are missing from the 10x10 multiplication table and why?

18. The numbers in the 2 row are called “even.” What would you call the numbers in the 5 row?

19. Can the same number appear more than once in a single column?

20. How many one syllable number names are found in a 10x10 multiplication table?

21. Think about dividing a number in the multiplication table by the one directly to its left. When does this result in a whole number?

22. Suppose you have as many different colored crayons as you want. Color in all 100 entries in a 10x10 multiplication table. Which numbers are the same color as 49 and why?

23. Choose a column or row and write a short story about the numbers in it.

24. Which row of the multiplication table is the hardest to remember and why?

25. If you were trying to fill in a blank multiplication table and got stuck, what would you do?

APPENDIX C

DIMENSIONS AND RATINGS FORMS

Dimension	Descriptive Definition
Creativity	The degree to which the math problem is creative, using your own subjective definition of creativity.
Pedagogical appropriateness	The degree to which the problem is appropriate for use in a classroom where the multiplication table has been learned.
Liking	How well you like the math problem, using your own subjective criteria for liking.
Clarity	The degree to which the math problem is written clearly.
Originality of idea	The degree to which the idea of the math problem is original.

Note: For all dimensions, each problem should be rated *relative* to the other 24 problems attached, and not in some absolute sense.

Problem 1	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 2	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 3	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 4	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 5	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 6	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 7	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 8	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 9	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 10	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 11	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 12	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 13	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 14	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 15	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 16	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 17	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 18	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 19	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 20	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 21	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 22	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 23	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 24	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Problem 25	Circle Exactly One Number For Each Dimension						
Creativity	(least)	1	2	3	4	5	(most)
Pedagogical appropriateness	(least)	1	2	3	4	5	(most)
Liking	(least)	1	2	3	4	5	(most)
Clarity	(least)	1	2	3	4	5	(most)
Originality of idea	(least)	1	2	3	4	5	(most)

Please check to ensure that all problems have been rated on each dimension.

Thank you very much for your participation!

APPENDIX D

IRB Approval

TEACHERS COLLEGE
COLUMBIA UNIVERSITY
OFFICE OF SPONSORED PROGRAMS

Institutional Review Board

February 21, 2014

Benjamin Dickman
1230 Amsterdam Avenue #1004
New York, NY 10027

Dear Benjamin,

Thank you for submitting your study entitled, *"Concepts of creativity in elementary school mathematical problem posing;"* the IRB has determined that your study is **Exempt** from committee review [Category **2**].

Please keep in mind that the IRB Committee must be contacted if there are any changes to your research protocol. The number assigned to your protocol is **14-178**. Feel free to contact the IRB Office [212-678-4105 or hersch@tc.edu] if you have any questions.

Please note that your consent form bears an official IRB authorization stamp. Copies of this form with the IRB stamp must be used for your research work.

Best wishes for your research work.

Sincerely,

Karen Froud, Ph.D.
Associate Professor of Speech and Language Pathology
Chair, IRB

cc: File, OSP

APPENDIX E

EMAIL CORRESPONDENCE

The main portion of the initial email read as follows:

> My name is Benjamin Dickman and I am in the process of completing my Ph.D dissertation in Mathematics Education at Columbia University. I am writing to you to ask whether you would be willing to participate in my study. Participation entails rating a total of 25 problems, posed by prospective mathematics teachers about the multiplication table, on five different scales.
>
> Feel free to decline, and note that even if you agree to participate, you can withdraw at any point should you not wish to proceed further. If you will be able to participate, then please notify me so that I can send along the corresponding items; similarly, if you will not be able to participate, then please let me know.

The main portion of the follow-up emails read as follows:

> Attached to this email are two documents: Dickman-Problems-Dimensions-Ratings and Dickman-IRB-Informed-Consent-Signature.
>
> The former document contains:
> 1. A list of twenty-five problems posed by prospective mathematics teachers about the multiplication table.
>
> 2. A list of the five dimensions with brief descriptive definitions.
>
> 3. Rating sheets to be filled out by participants.
>
> The latter document is the IRB form that must be signed for collected data to be used.
>
> There are four different ways to return the filled-out documents:
> 1. Attaching them to an email.
>
> 2. Faxing them to the Teachers College program in Mathematics Education.
>
> 3. Sending by snail mail.
>
> 4. Emailing back with an address to which I can send a self-stamped envelope.
>
> Thank you for your consideration in participating in my study!