

Running Head: Situative Creativity

Situative Creativity: Larger Physical Spaces Facilitate
Thinking of Novel Uses for Everyday Objects

Joel Chan^a and Timothy J. Nokes-Malach^b

^aHuman-Computer Interaction Institute, Carnegie Mellon University, 2504B Newell-Simon Hall,
5000 Forbes Ave, Pittsburgh, PA 15213, joelchuc@cs.cmu.edu, 479-647-0575 (tel),
corresponding author

^bLearning Research and Development Center, University of Pittsburgh

ABSTRACT

People often use spatial metaphors (e.g., think “laterally”, “outside the box”) to describe exploration of the problem space during creative problem solving. In this paper, we probe the potential cognitive underpinnings of these spatial metaphors. Drawing on theories of situative cognition, semantic foraging theory, and environmental psychology, we formulate and test the hypothesis that larger physical spaces can facilitate divergent (but not convergent) processes in problem space exploration. Across two experiments, participants worked on a battery of problem solving tasks intended to represent divergent (alternative uses, shape invention) and convergent (remote associates, letter extrapolation) problem solving processes in either a large or a small room. In Experiment 1, participants in the larger room produced more novel alternative uses for everyday objects, and created more novel shape inventions, but generated less practical alternative uses, than participants in the smaller room. In Experiment 2, participants in the larger room (including a variant larger room) also produced more novel alternative uses for everyday objects, and less practical alternative uses, than participants in a small room, but did not create more novel shape inventions. These results suggest that spatial metaphors for problem space exploration may reflect meaningful cognitive phenomena: people may be able to search more broadly in a problem space if they are in an environment where broad physical search is a salient affordance; however, this effect appears to be relatively small and may depend on having sufficiently motivated participants.

INTRODUCTION

A key component of creative problem solving is exploration of the problem space. The *problem space* is typically described as the mental representation of the problem, including the initial problem description, goal, and operators (i.e., strategies) to move from the initial state to the goal state (Newell & Simon, 1972). Theories of creative problem solving posit that the effective initial exploration of the problem space — sometimes called “divergence” or “divergent thinking” (Guilford, 1956) — is critical to produce a successful solution (Amabile, 1983; Finke, Ward, & Smith, 1996; Sawyer, 2012; Simonton, 2011; Wallas, 1926; Warr & O’Neill, 2005). Effective problem space exploration can be supported by considering many different solution approaches (Adánez, 2005; Chan et al., 2011; Parnes & Meadow, 1959; Shah, Millsap, Woodward, & Smith, 2012; Torrance, 1988), increasing the variance in the quality of solutions considered (e.g., being willing to consider “wild” ideas; Chan et al., 2011; Girotra, Terwiesch, & Ulrich, 2010; Terwiesch & Ulrich, 2009), considering solutions and perspectives from outside one’s discipline or problem domain (Chan et al., 2011; Gentner & Markman, 1997; Ward, 1998), relaxing inferred constraints about the problem description (Knoblich, Ohlsson, Haider, & Rhenius, 1999; Ohlsson, 1992), and exploring alternative conceptualizations of the problem (Kaplan & Simon, 1990; MacGregor & Cunningham, 2009). Divergence can also be facilitated by modulation of attention: for example, a reduction of attentional control or focus has been identified as a key mechanism for achieving divergent thinking and making remote associations in creative problem solving (Aiello, Jarosz, Cushen, & Wiley, 2012; Haarmann, George, Smaliy, & Dien, 2012; Martindale, 1997; Wiley & Jarosz, 2012).

Successful exploration is often described with spatial language and imagery. For example, people commonly encourage one another to think “laterally”, not “vertically” (Bono,

1970), “outside the box”, or explore “broadly” (Wiley, 1998), and make “remote” associations in semantic memory (Mednick, 1962). In this paper, our goal is to probe the potential cognitive underpinnings of these spatial metaphors. Are these metaphors arbitrary, or merely artifacts of human convention? Or do they identify real cognitive phenomena? Could embodying variations in these spatial metaphors (e.g., large vs. small physical environments) influence the nature of people’s search patterns in semantic space?

The present investigation is inspired by a growing body of literature across a diverse range of tasks that suggests that people’s embodied physical context can have significant implications for information processing. For example, people perceive slopes as steeper if they wear a heavy vs. a light backpack (Bhalla & Proffitt, 1999), can transfer knowledge and skills across contexts when there is high interconnectedness across activities and practices in those contexts (Brown, Collins, & Duguid, 1989), and make more “holistic” decisions (i.e., integrating multiple sources of data and abstraction) in rooms with higher vs. lower ceilings (Meyers-Levy & Zhu, 2007).

The literature provides two potential theoretical motivations for suspecting that the spatial metaphors of creative search have grounding in cognitive phenomena. The first account, which we call “direct priming”, is exemplified by Hills and colleagues’ argument that goal-directed search for resources in external spaces and search for resources in internal spaces (e.g., semantic memory) share a common neural substrate (Hills, 2006; Hills, Jones, & Todd, 2012; Hills, Todd, & Goldstone, 2008). Specifically, that they share dopaminergic modulation of area-restricted search such that search is narrow in situations where the target resources have been frequently found in the past and search becomes broad in situations where the target resources are encountered less frequently. One intriguing implication of this argument is that expectations

about the structure of search environments in external spaces can shape search patterns in internal spaces, or vice versa.

To test this implication, Hills et al. (2008) studied how search patterns on an anagram task (i.e., search for as many words as possible for a given letter set) might be shaped by prior experience with a spatial foraging task (i.e., search for high-value pixels in a simple 2D computer maze). In their experiment, participants completed spatial foraging tasks with either “clumpy” — many pixels concentrated in a few patches — or “diffuse” — pixels evenly distributed in the environment — resource distributions, and subsequently tried to find as many anagrams as they could. They found that participants who had just experienced a “clumpy” distribution of pixels took longer to switch between letter sets when searching for anagrams, consistent with expectations for a “clumpy” distribution of anagrams (i.e., expecting letter sets to contain more anagrams). They inferred from this that the distribution of resources in the spatial environment primed expectations for the distribution of “resources” in the semantic space. This analysis suggests that the affordances in the external, physical environment (e.g., the distribution of resources) may shape the mind’s internal search in semantic space.

Following this line of thought, we reason that, to the extent that large physical environments afford free movement and exploration, they may also better facilitate divergent problem solving (i.e., exploration of semantic space) relative to smaller, constrained spaces. Rather than simply fostering increased performance via increased effort, people might be sensitive to how larger physical spaces afford freer exploration, and consequently adopt a semantic search strategy that better matches this resource distribution, such as by relaxing their focus of attention from more clearly relevant or high-quality responses to more semantically distant and varied (and likely more novel) associations (Aiello et al., 2012; Haarmann et al.,

2012; Martindale, 1997; Wiley & Jarosz, 2012). This direct priming of attentional focus might occur without conscious awareness, similar to some varieties of top-down modulation of visual attention (Koch & Tsuchiya, 2007). This direct priming mechanism can also be related to the notion of “frames” in research on situative cognition: expectations (whether explicit or tacit) about a given situation that are influenced by the affordances and constraints of particular environments, and go to shape cognition and interaction (Goffman, 1974; Greeno, 1994; Greeno & Middle-School Mathematics through Applications Project Group, 1998; Maclachlan & Reid, 1994; Nokes-Malach & Mestre, 2013; Scherr & Hammer, 2009). Because both varieties of direct priming mechanisms can occur without conscious awareness, we do not expect facilitation of divergent performance to be associated with more effortful performance. Indeed, to the extent that people relax their focus of attention to search more broadly we might even expect to see a *decreased* perception of task difficulty as measured by cognitive load (e.g., Antonenko, Paas, Grabner, & Gog, 2010; Chandler & Sweller, 1991).

The second line of reasoning, which we call the “concept activation” account, comes from research in environmental psychology that explores how certain configurations of physical environments can prime certain psychological states or ideas, which can then influence later information processing. For example, Hall (1966) argues that small and contained spaces (e.g., chapels) can evoke notions of confinement or restrictedness, while larger spaces (e.g., cathedrals) can prime notions of freedom and openness. Similarly, Moore and colleagues (1994) suggest that lower ceilings may invoke more restricted play, while higher ceilings may encourage “freer” play. In the Meyers-Levy and Zhu (2007) study mentioned previously, the effect of the ceiling height manipulation on decision-making was mediated by activation of the concept of “freedom” vs. “confinement”. This line of reasoning presents a indirect mechanism by which larger spaces

prime concepts of “freedom” and “broadness”, which in turn induces information processing that is also “less constrained” (e.g., more holistic, as in Meyers-Levy & Zhu (2007), thereby facilitating divergent processing during problem solving.

In contrast to the direct priming account, concept activation may also be marked by affective changes (e.g., increases in positive affect, decreases in negative affect), since concepts related to “freedom” may have positive valence, while concepts related to “confinement” may have negative valence. For example, a recent affective norming project found that the word “freedom” had a highly positive valence score of 7.72 on a 1 to 9 valence scale (1 is highly negative, 9 is highly positive), while the word “restrict” had a much more negative valence of 3.48 out of 9 (Warriner, Kuperman, & Brysbaert, 2013). Therefore, measuring changes in affect may be a way to distinguish between direct priming or concept activation accounts of potential associations between physical surroundings and divergent/convergent problem solving processes: increased divergent performance in large physical spaces accompanied by increases in positive affect (and decreases in negative affect), would be more consistent with a concept activation account of the cognitive basis of spatial imagery for divergent exploration.

Synthesizing these ideas, we test the hypothesis that larger spaces will have a facilitation effect on divergent problem-solving processes (i.e., processes that have similar cognitive characteristics to the exploration stage of the creative process), but not “convergent” problem solving processes (i.e., processes that focus on “converging” on a single “correct” or canonical answer). To the extent that increased divergence may be at odds with convergent processes (Goldenberg, Larson, & Wiley, 2013), larger spaces might also *hinder* convergent problem solving. We further hypothesize that this facilitation would be accompanied by decreases in perceived task difficulty. Measures of affect might help distinguish between the direct priming

and concept activation explanations of observed effects. In summary, in this paper, we examine the following two main hypotheses:

H1: Divergent performance will be higher in larger vs. smaller physical spaces.

H2: Convergent performance will be lower in larger vs. smaller physical spaces.

H3: Perceived task difficulty for divergent tasks will be lower in larger vs. smaller physical spaces.

We conducted two experiments to test these hypotheses, first with a sample of paid volunteers, and then with a larger sample of psychology subject pool participants and an expansion of the range of physical spaces and problem solving stimuli that are tested. To preview our results, we find partial support for the first two hypotheses across both experiments, and find that these effects are not associated with changes in affect.

EXPERIMENT 1

In this study, we provide a first test of the three hypotheses. The basic experimental approach is to have participants work on a battery of problem solving tasks intended to represent both divergent and convergent processing in either large or small rooms. As noted, our hypotheses are that *divergent* problem solving performance will be facilitated by being in a large (vs. small) room, while convergent performance will be hindered by being in the large room.

Methods

Participants

Forty-seven people (20 males, 27 females; ages 19-66, average age 27) from the community at a large research university in the Northeast United States participated in this study. Thirty-five of the participants were undergraduate or graduate students at the university. Most of the other participants were recent graduates or employees of the university or businesses on

campus. All participants were recruited through fliers posted around campus and were compensated 10 dollars for their time.

Four participants didn't produce valid data on one of the problem solving tasks (3 did not produce any valid inventions, and 1 did not produce any valid uses), and were therefore dropped from our analyses: there were two each from the large and small rooms respectively¹. Therefore, our final dataset consisted of 43 participants.

Materials

Room Size Manipulation. To manipulate room size, we had participants complete their problem solving tasks in one of two rooms on campus. The “large” room was a conference auditorium (see Fig 1, left panel). The dimensions of the room were approximately 15' W x 30' L



Figure 1. Picture of large room (left) and small room (right) in Experiment 1.

¹ Results are the same with partial data from these participants included

x 15' H. Participants completed their tasks on a desk in the front of the auditorium facing towards the audience seats so that the size of the room would be salient. Other than the desk and chair the participants used, and the other chairs facing the front of the auditorium, the auditorium was empty.

The “small” room was a former office space that was emptied out for the experiment (Fig 1, right panel). The dimensions of the room were approximately 8' W x 10' L x 8' H. Participants completed their tasks on a desk facing one of the walls. It was empty except for the desk and chair the participants used.

Other than the size of the room, we made sure that the two rooms were similar in a number of important ways, including amount of stimuli encountered on the walk to the room (both rooms were in the same building), ambient noise (we chose rooms that were far from other offices in the building), and temperature (both rooms shared the same central air conditioning system). The one potentially salient difference was the tone of lighting: the large room used incandescent lighting, while the small room used fluorescent lighting.

Problem-Solving Tasks

Participants completed a battery of four problem-solving tasks intended to represent both divergent and convergent processing: 1) an alternative uses task, 2) a shape invention task, 3) a version of the Remote Associates Test (RAT), and 4) a letter series extrapolation task. In this section we describe each task along with the hypothesized processes involved.

Alternative Uses. The alternative uses task is patterned after Guilford's (1967) classic alternate uses task in which the problem solver is asked to list as many uses as possible for a common object (e.g., think of as many uses as you can for a brick). It has been hypothesized to measure divergent thinking processes because the output of the task is a range of responses rather

than one correct or incorrect response. Task output is typically measured in terms of the fluency and flexibility (e.g., novelty) of the responses. However, this task may also involve convergent processes. People may initially think of a variety of responses, and subsequently evaluate and select only the uses that are both novel *and* practical. This corresponds to the selection/evaluation processes/phases in various theories of cognitive and creative production, such as the convergent production component of Guilford's (1956) "structure of intellect" theory, the response validation stage of Amabile's (1983) process model of creativity, and the Explore phase of Finke and colleagues' (1996) Geneplore model. Therefore, we hypothesize that fluency and novelty are measures of divergent thinking on this task, whereas practicality is a measure of convergent thinking. It is useful to note that convergence may not necessarily always follow divergence: convergence can also reflect attentional focus on and rapid selection of the most readily accessible responses, which are often the most successful or appropriate (Bilalić, McLeod, & Gobet, 2008; Guilford, 1956; Luchins, 1942).

We used "SHOE" and "NEWSPAPER" as our common object items and gave the following instructions to participants: *"In this part of the experiment, your task is to list as many uses as you can for an object (named below). For example, if the object is "BRICK", you could say "building material, doorstep, anchor, etc.". The goal is to come up with as many uses of an object as possible. There are 2 of these problems, and you will have 4 minutes for each."*

Shape Invention. In the shape invention task (Finke et al., 1996), the problem solver is given three three-dimensional shapes to combine together to create as many useful objects that belong to one of three given categories (e.g., toys and games, transportation). Similar to the alternative uses task, we hypothesize that this task includes elements of both divergent and convergent processing. Again, we hypothesize that the novelty of the items generated would

reflect divergent thinking processes, whereas the practicality of the objects would reflect convergent processes.

Some versions of this task allow the problem solver to choose the shapes and/or the object categories; we elected to randomly select both shapes and categories and present them as givens for all participants. Part of the motivation for this was that Finke et al. (1996) found that this condition stimulated the most creative responses, and we wanted to give our participants the best chance of displaying creative performance. Finke et al.'s original set of shapes includes 15 different shapes, ranging from spheres to cylinders, to wires, wheels, and flat squares; each shape belongs to a sub-category set of 5 shapes, ordered by how difficult it is (normed from their studies) to incorporate into an invented object (easy, medium, and hard). We randomly sampled 3 shapes from this list, with “easy” shapes having a probability = 0.10 of being selected, “medium” shapes having a probability = .07 of being selected, and “hard” shapes having a probability = 0.03 of being selected. Finke et al.’s original set of object categories consisted of 8 categories: weapons, toys and games, appliances, transportation, scientific instruments, tools and utensils, furniture, and personal items. We randomly selected three categories from this list to give to participants. We ended up with the sphere, tube, and cone objects (see Fig 2), and the following object categories: tools and utensils, toys and games, and personal items.

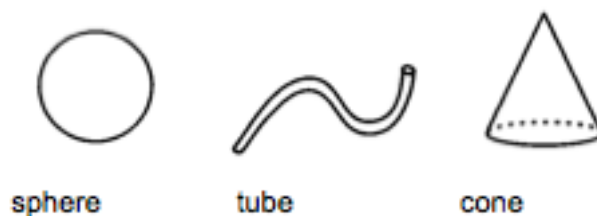


Figure 2. Shapes for invention task in Experiment 1.

We gave the following instructions to participants: *“In this part of the experiment, your task is to try to use the following “parts” to “construct” as many useful objects as you can. These objects can be existing things or things you invent. You will have 8 minutes to do this. The rules for using the parts to construct objects are as follows: 1) you are allowed to vary the size, position, or orientation of any part, but you may not bend or deform the parts (except the tube), 2) the parts can be put inside one another, 3) you decide if the parts are hollow or solid, and 4) you decide what material the parts are made of – they can be made of any material, including wood, metal, plastic, rubber, or glass, or any combination of these materials.”*

RAT. To complete the RAT (Bowers, Regehr, Balthazard, & Parker, 1990; Mednick, 1962; Mednick & Mednick, 1967), solvers must generate a target word that is related to a list of three cue words. For example, a correct response for the cue words Elephant-Lapse-Vivid would be the target word “memory.” While the RAT is most commonly used as a measure of creativity (e.g., Wiley, 1998), and several prior studies have examined the role of divergent processes in RAT performance (Aiello et al., 2012; Haarmann et al., 2012), we reasoned that it also heavily taps convergent processes since the final output is a *single answer* that is compared to a predetermined correct answer. We are not the first to treat the RAT in this way: other recent studies have also studied the RAT as primarily a convergent task in contrast to the alternate uses task (treated as primarily measuring divergent processes), and found both dissociable effects on these tasks from their manipulations and a lack of correlation between performance on these tasks (Colzato, Ozturk, & Hommel, 2012; Oppezzo & Schwartz, 2014; Radel, Davranche, Fournier, & Dietrich, 2015). We gave participants 32 items, of varying difficulty, drawn from Mednick and Mednick (1967) and Bowers, Regehr, Balthazard, and Parker (1990), chosen to reflect a range of difficulty levels. The full list of items can be seen in the appendix.

Letter Series. In the letter series task, the problem solver is given a series of letters and is asked to generate additional letters to “N” places to complete the pattern exemplified in the given series. For example, given the series “aaabbbccc”, the correct extrapolation to N=3 places is “ddd”. Similar to the RAT this task requires the identification and generation of a single answer. Although participants may consider several possible patterns when solving the problem they must eventually converge or decide on one to extrapolate. We gave participants 18 items drawn from prior studies with the letter series task (Nokes, 2009; Simon & Kotovsky, 1963), which were chosen to reflect a range of difficulty levels. The items had initial series ranging from 9 to 16 letters in length; all items required participants to extrapolate the series to N=10 places. The full list of items can be seen in the appendix.

Dependent Measures

In the previous section we described each task we used for the experiment. Some of the tasks (specifically the uses and invention tasks) were hypothesized to include both divergent and convergent processing. We now describe how we measured divergent and convergent processing across the tasks.

Divergent Measures. Both the uses and invention tasks were scored for fluency and novelty to yield our primary divergent measures. Fluency was defined as the number of uses or inventions generated. Novelty was rated on a scale from 1 (not at all novel) to 4 (extremely novel). Examples of low and high novelty uses are “use SHOE to protect feet” and “use SHOE as boat for termites”; examples of low and high novelty inventions are shown in Figure 3. Two trained coders evaluated the uses, with high inter-rater reliability, $ICC(2,2) = .89$. Three trained coders evaluated the inventions, with high inter-rater reliability, $ICC(2,3) = .85$. Each use/invention’s novelty score was the arithmetic mean of all judges’ scores for that use/invention.

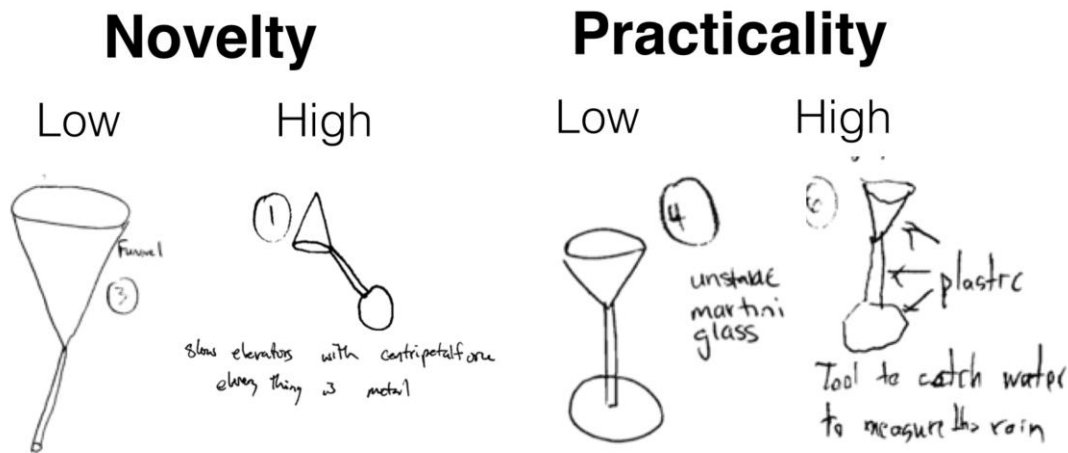


Figure 3. Example low and high novelty/practicality inventions. The low and high novelty inventions are a “funnel” and a device that “slows elevators with centripetal force”. The low and high practicality inventions are an “unstable martini glass” and a “tool to catch water to measure the rain”. Note that the shapes used for invention are a cone, sphere, and tube.

Novelty scores were then aggregated into participant-level measures in the following ways: *mean novelty* (how novel their uses were, on average), and *max novelty* (what was the highest novelty score they achieved).

Convergent Measures. Both the uses and invention tasks were also scored for practicality, to reflect convergent processing on those tasks. Practicality scoring for the uses task was initially done with a 4-point scale (1 - unlikely to work at all, 2 - will work less well than conventional means, 3 - will work as well as conventional means, 4 - will work better than conventional means), but was collapsed to a 3-point scale because there were almost no (agreed-upon) generated uses that warranted a 4. Two trained coders evaluated the uses, with good inter-rater reliability, $ICC(2,2) = .79$. In contrast, scoring for the invention category had slightly higher variance, allowing us to code inventions on a 5-point scale (1 – extremely bad example of its invention category, to 5 – exceeds expectations for a good example of its invention category). Seven trained coders evaluated the inventions for practicality, with acceptable inter-rater

reliability, $ICC(2,7) = .61$. Examples of low and high practicality uses are “use SHOE as tent stake” and “use SHOE as slapping device to bring someone back to their senses”; examples of low and high practicality inventions are shown in Figure 3. Participant-level practicality measures were created by taking the arithmetic mean of practicality scores achieved to create a *mean practicality* measure for both uses and invention tasks, separately.

Performance on the RAT and letter series tasks were intended to primarily reflect convergent processing, since both tasks sought the production of a single “best” response. One trained coder scored the RAT responses as either correct or incorrect, using the answer key from Mednick & Mednick (1967) and Bowers et al., (1990). Percent correct was used for analysis. Letters series task performance was measured by marking responses as either correct or incorrect, using canonical answers from the prior references (Nokes, 2009; Simon & Kotovsky, 1963), and the percent correct was used for analysis.

Other Measures

Perceived Task Difficulty. To measure perceived task difficulty, we adapted two items from prior research with cognitive load (Jang & Schunn, 2012). The measure was about the task just completed. The first item asked, “*how easy or difficult was this task?*”, and participants were asked to answer using a 1 to 9 scale, where 1 was anchored as “very, very easy”, and 9 was anchored as “very, very difficult”. The second item asked “*how much mental effort (e.g., searching, remembering, thinking, deciding) did the task take?*”, and participants were asked to answer using a 1 to 9 scale, with 1 anchored as “very, very low mental effort”, and 9 anchored as “very, very high mental effort”.

Positive and Negative Affect. To measure affect, we used the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988), in which a subject is given 20 words that

describe different feelings and emotions, and — using a scale of 1 (very slightly or not at all) to 5 (extremely) — rates the extent to which she feels that way “right now” (i.e., at the present moment).

Procedure

Participants were greeted and brought to either the large or small room, depending on their assignment. They were then informed that they would complete four problem-solving tasks, grouped into two blocks with two problem sets each. The uses and invention tasks formed one block, while the RAT and letter series task formed the other block. The tasks were counterbalanced by block, and specific problem solving tasks within the block (RAT and letters; uses and invention). They had eight minutes to complete each problem set, and a one-minute warning was given before the time was up.

After each problem set, participants completed the untimed cognitive load survey. Before they began the problem sets, participants were asked to complete the first PANAS to get a baseline measure. They were asked to complete a second and third PANAS after the first and second blocks of problem sets, respectively. Overall, the experiment ran no longer than 45 minutes. The experimenter remained in the room during the length of the experiment, seated behind the participant.

Design

This study had a between-subjects design. The independent variable was room size, with two levels (large or small). Participants were randomly assigned to conditions. In the final dataset, there were 21 participants in the large room and 22 participants in the small room.

Results

Table 1 shows the descriptive statistics for all problem solving measures. Table 2 shows

the intercorrelations between participants' problem solving task performance measures, collapsed across Experiments 1 and 2. In general, the correlations were in the expected directions (e.g.,

Table 1

Descriptive Statistics for Experiment 1

	Mean	Median	Min	Max	SE
Divergent measures					
Uses fluency	22.07	21	7	42	1.32
Uses novelty mean	1.67	1.59	1	2.58	0.06
Uses novelty max	3.12	3	1	4	0.13
Invention fluency	5.07	5	2	9	0.29
Invention novelty mean	2.74	2.83	1.56	3.83	0.08
Invention novelty max	3.58	3.67	2	4	0.08
Convergent measures					
Uses practicality	2.62	2.64	2.03	3	0.04
Invention practicality	3.15	3.21	1.79	3.86	0.06
RAT	0.59	0.62	0	1	0.04
Letters	0.61	0.71	0	1	0.05

Table 2

Intercorrelations between measures, collapsed across Experiments 1 and 2

	Divergent					Convergent			
	Uses novel mean	Uses novel max	Invent fluen.	Invent novel mean	Invent novel max	Uses pract.	Invent pract.	RAT	Letters
U fluency	0.50*	0.50*	0.37*	0.24*	0.26*	-0.43*	-0.07	-0.04	-0.01
U novel mean		0.79*	0.31*	0.27*	0.29*	-0.89*	-0.02	0.01	-0.04
U novel max			0.17*	0.25*	0.22*	-0.66*	0.01	0.03	-0.12
I fluency				0.09	0.46*	-0.32*	0.00	-0.07	-0.06
I novel mean					0.75*	-0.23*	-0.11	-0.02	-0.03
I novel max						-0.28*	-0.05	-0.01	-0.02
U practicality							0.06	-0.06	0.00
I practicality								0.14 ^m	0.00
RAT									0.21*

^m $p < .10$; * $p < .05$

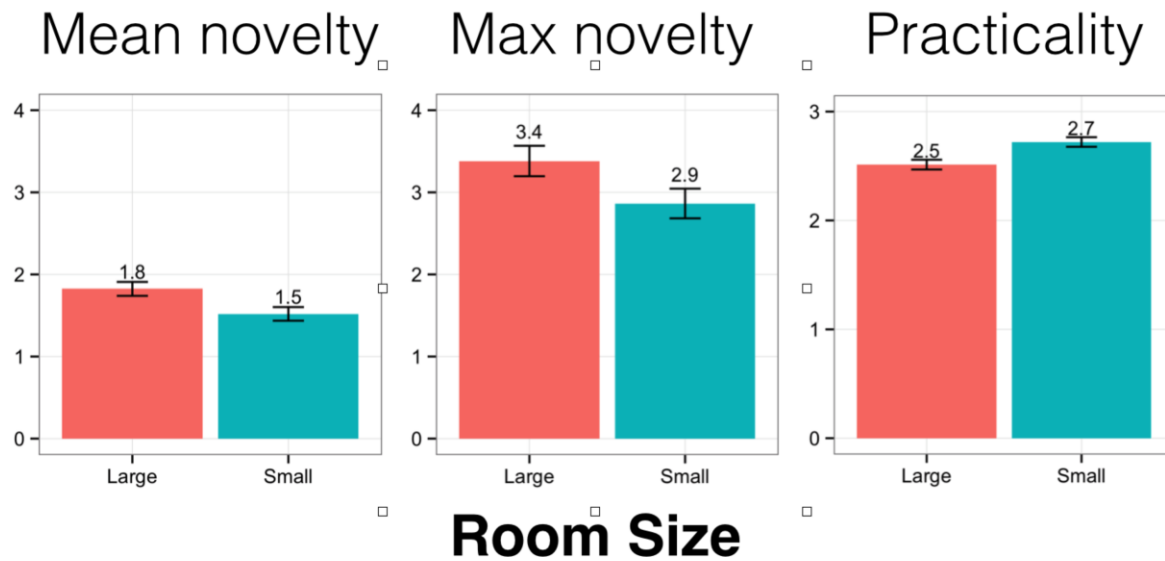


Figure 4. Novelty and practicality of alternative uses by room size in Experiment 1. Error bars are ± 1 SE.

significant positive correlations between divergent measures, significant positive correlations between letter series and RAT measures, significant negative correlations between uses practicality and the divergent measures). However, the correlations are generally low, explaining small amounts of common variance, and some correlations are missing (e.g., no significant correlations between practicality and the letters series and RAT measures, and no significant correlations between invention practicality and any of the other measures). This suggests that the measures do not necessarily primarily reflect the two constructs of divergence and convergence (as we had hypothesized). Therefore, we analyze each measure separately.

Alternative Uses

There was an effect of room size on mean fluency, with participants in the large room generating more uses ($M=25.0$, $SE=1.8$) than participants in the small room ($M=19.3$, $SE=1.8$), $d=0.68$, 95% CI=[0.03, 1.33], $F(1,41)=5.0$, $p=.03$. There was also an effect of room size on mean novelty, with participants in the large room producing higher mean novelty with their uses

($M=1.8$, $SE=.09$) than those in the small room ($M=1.5$, $SE=.08$), $d=0.78$ [0.13, 1.44], $F(1,41) = 6.6$, $p=.01$ (Fig. 4, left panel). The results were similar for max novelty: participants in the large room achieved marginally higher max novelty scores ($M=3.4$, $SE=.19$) than those in the small room ($M=2.9$, $SE=.18$), $d=0.61$ [-0.04, 1.26], $F(1,41)=4.0$, $p=.05$ (Fig. 4, middle panel).

In contrast, participants in the large room generated alternative uses that were significantly *less* practical ($M=2.5$, $SE=.04$) than those from the small room ($M=2.7$, $SE=.04$), $d=-1.00$ [-1.68, -0.33], $F(1,41)=10.9$, $p=.00$ (Fig. 4, right panel).

Shape Invention

Participants in the large room generated slightly more inventions ($M=5.5$, $SE=.42$) than participants in the small room ($M=4.6$, $SE=.41$), $d=0.47$ [-0.17, 1.11], but this difference was not statistically significant, $F(1,41)=2.3$, $p=.13$. However, there was an effect of room size on mean novelty, with participants in the large room, on average, creating more novel inventions ($M=2.9$, $SE=.10$) than participants in the small room ($M=2.6$, $SE=.10$), $d=0.76$ [0.10, 1.41], $F(1,41)=6.2$, $p=.02$. Similarly, there was an effect of max novelty, with the most novel inventions of participants in the large room being, on average, more novel ($M=3.8$, $SE=.10$) than the most novel inventions of participants in the small room ($M=3.4$, $SE=.10$), $d=0.76$ [0.10, 1.41], $F(1,41)=6.2$, $p=.02$. For practicality, there was no effect of room size, with inventions from the large room condition about as practical ($M=3.2$, $SE=.08$) as those from the small room ($M=3.1$, $SE=.08$), $d=0.32$ [-0.31, 0.96], $F(1,41)=1.1$, $p=.29$.

RAT

There was no effect of room size, with participants in the large room having about the same mean proportion correct ($M=.57$, $SE=.05$) as participants in the small room ($M=.60$, $SE=.05$), $d=-0.12$ [-0.75, 0.51], $F(1,41) = 0.2$, $p=.69$.

Letter Series

Participants in the large room generated a slightly higher proportion of correct responses ($M=.65$, $SE=.07$) than participants in the small room ($M=.57$, $SE=.07$), $d=0.22$ $[-0.41, 0.85]$, but this difference was not statistically significant, $F(1,41) = .05$, $p=.48$.

Perceived Task Difficulty

There was a marginal effect of room size on perceived difficulty for the uses task, with participants in the large room self-reporting slightly lower levels of difficulty ($M=4.1$, $SE=.36$) compared to participants in the small room ($M=5.1$, $SE=.36$), $d=-0.53$ $[-1.15, 0.08]$, $F(1,45)=3.4$, $p=.07$. For the invention task, there was a main effect of room size, with lower self-reported difficulty in the large ($M=5.2$, $SE=0.3$) vs. small room ($M=6.8$, $SE=0.3$), $d=-1.16$ $[-1.81, -0.51]$, $F(1,45)=15.7$, $p=0.00$.

In contrast, for the RAT, participants in the large room self-reported about the same levels of perceived difficulty ($M=7.1$, $SE=.30$) as participants in the small room ($M=7.4$, $SE=.29$), $d=-0.16$ $[-0.77, 0.43]$, $F(1,45)=0.34$, $p=.56$. Similarly, for the letters task, participants in the large room self-reported about the same levels of perceived difficulty ($M=6.3$, $SE=.36$) as participants in the small room ($M=5.9$, $SE=.36$), $d=0.24$ $[-0.37, 0.84]$, $F(1,45)=0.65$, $p=.42$.

Positive and Negative Affect

There was no effect of room size on positive affect, with participants in the large room self-reporting about the same levels of positive affect ($M=27.6$, $SE=1.5$) as participants in the small room ($M=27.4$, $SE=1.5$), $d=0.04$ $[-0.59, 0.68]$, $F(1,41)=0.02$, $p=.89$. Similarly, there was no effect of room size on negative affect, with participants in the large room self-reporting about the same levels of negative affect ($M=12.3$, $SE=.47$) as participants in the small room ($M=11.8$, $SE=.46$), $d=0.22$ $[-0.41, 0.86]$, $F(1,41)=0.53$, $p=.47$.

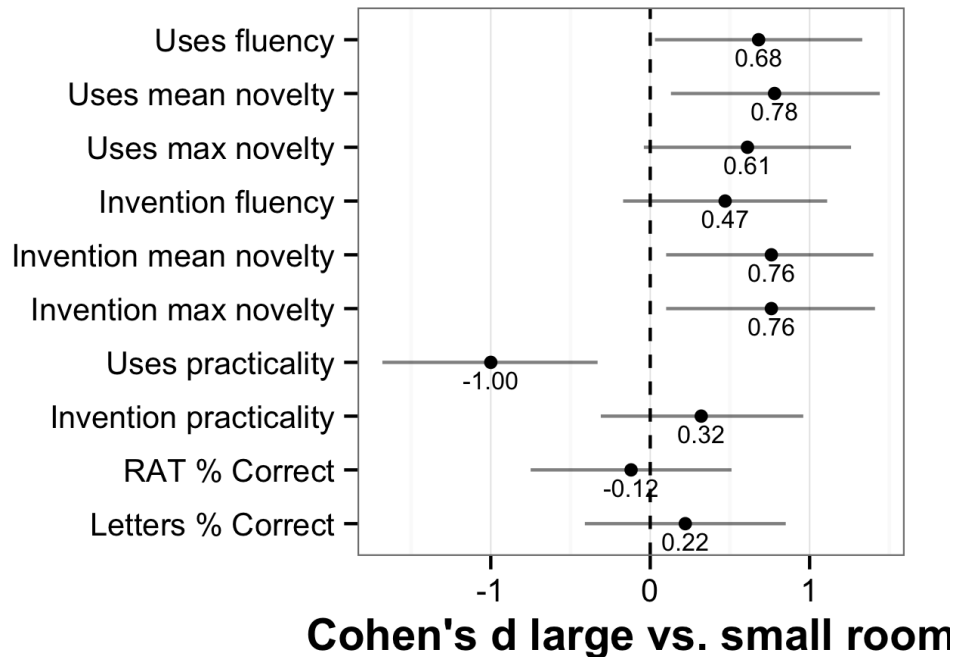


Figure 5. Summary of effects in Experiment 1. Error bars are 95% CIs.

Discussion

Experiment 1 yielded evidence consistent with Hypothesis 1 (see Figure 5 for a summary of the observed effects). As predicted, participants' performance was higher on the divergent problem solving measures in the larger room than in the smaller room (e.g., uses fluency, uses novelty, invention novelty). In contrast, we found only partial support for Hypothesis 2, i.e., that participants in the large rooms would perform worse on convergence measures than participants in the small rooms. Consistent with the hypothesis, we found that participants in the larger room showed lower performance on the uses practicality measure. However, no differences were found across the two groups on invention practicality and performance on the RAT and letter series task. In the general discussion we discuss possible reasons for why we did not observe stronger negative effects of a large space on convergence measures. Taken together, these results show that the benefits of the larger room for divergent performance were not simply due to a general facilitation effect of being in the larger room: rather, there seems to be a specific effect of being

in a larger room on the cognitive processes that enable divergent performance.

Analysis of the additional measures yielded additional insights. In support of Hypotheses 3, the perceived task difficulty results suggest that participants in the larger room not only performed better on the uses and invention tasks (in terms of divergent performance measures), but also found the task overall to be less cognitively taxing (compared to participants in the smaller rooms), suggesting that some of the performance benefits might be due to unconscious mechanisms (e.g., automatic attunement of semantic search patterns to search affordances in the physical environment). Further, analysis of the survey responses for PANAS suggest that the differences are not explained by positive boosts to affect in the larger room (or increased negative affect in the smaller room). This result is consistent with the direct priming hypothesis and not the concept activation account.

EXPERIMENT 2: REPLICATION AND EXTENSION

Given the novelty of our hypotheses, we conducted a second study to replicate and extend the results of Experiment 1. The focus of the extension is to ensure that the effects were not due to idiosyncrasies of the particular configurations of the large room or problem solving stimuli. To this end, we slightly altered the large room manipulation from Experiment 1 (participants sat at the top of the auditorium rather than at the bottom), and added a second new large room that had a lower ceiling height but was still spacious horizontally. To maximize statistical power, we treated them as a single condition in our analyses². We also changed the objects used for the uses task, as well as the categories and shapes used for the invention task.

² There were no statistical differences between participants in the two large rooms on any of the measures.



Figure 6. Picture of new large room in Experiment 2.

Methods

Participants

One hundred and nine undergraduates (61 females; ages 18-31, average age 19) enrolled in Introduction to Psychology at a large research university in the Northeast United States participated in this study. All participants were recruited through the university's psychology subject pool, and were compensated with course credit.

Three participants (assigned to the large rooms) did not produce any valid responses to the invention task, and were dropped from all analyses, leaving us with 106 total participants in our final dataset³.

Materials

Room Size Manipulation. In this study, the “large” setting of our room size manipulation included two different rooms: 1) the same conference auditorium as in Experiment 1, with the only difference being that participants sat at the top of the room, rather than the bottom of the

³ Results are the same with partial data from these participants included

room, and 2) another conference room in the same building (see Fig. 6). The dimensions of the new large room were approximately 15' W x 30' L x 8' H. Other than the desk and chair the participants used, along with the other desks and chairs in the room, the room was empty. Note also that the lighting here is fluorescent, similar to the small rooms in both experiments.

The “small” room was another former office space in the same building that was emptied out for the experiment. The dimensions of the room were the same as in Experiment 1, i.e., approximately 8' W x 10' L x 8' H. As in Experiment 1, participants completed their tasks on a desk facing one of the walls. It was empty except for the desk and chair the participants used.

Problem-Solving Tasks. As in Experiment 1, participants completed the alternative uses, shape invention, RAT, and letter series tasks. The only differences from Experiment 1 are with the stimuli for the alternative uses and shape invention tasks. The objects used for the alternative uses were “CUP” and “TABLE”. A different set of categories and shapes were randomly sampled for the invention task, using the same procedure as in Experiment 1. The new categories were “Transportation”, “Furniture”, and “Weapons”, and the new shapes were “rectangular block”, “ring”, and “half-sphere” (see Fig. 7).

Dependent Measures

As in Experiment 1, we obtained measures of fluency and novelty for both the uses and



Figure 7. Shapes for invention task in Experiment 2.

invention tasks. Inter-rater reliability was high for novelty scoring across both tasks, $ICC(2,2) = .83$ for uses novelty, and $ICC(2,3) = .84$ for invention novelty. We then aggregated scores at the participant-level into mean and max novelty. Also as in Experiment 1, we also evaluated uses and inventions for practicality. Inter-rater reliability was high for uses practicality, $ICC(2,2) = .82$, and acceptable for invention practicality, $ICC(2,4) = .67$. We then aggregated scores into participant-level measures of mean uses and invention practicality. Both the RAT and letter series tasks were scored identically to Experiment 1 (i.e., percent correct).

Procedure

The procedure was identical to Experiment 1, except that participants did not complete the PANAS measurement at any point.

Design

This experiment had a between-subjects design. The independent variable was room size, with two levels (large or small). Participants were randomly assigned to conditions. In the final dataset, there were 68 participants in the large rooms, and 38 participants in the small room.

Results

Descriptive statistics for all measures are shown in Table 3. Note that performance on all divergent measures (except for uses fluency) was significantly lower than that observed in Experiment 1. In contrast, performance on the practicality measures for both the uses and invention tasks were significantly higher than in Experiment 1. However, letters performance was also lower than in Experiment 1. We return to this issue in the discussion when interpreting the relationship between the results across the two experiments.

Table 3

Descriptive Statistics for Experiment 2

	Mean	Median	Min	Max	SE
Uses fluency	19.93	19	6	37	0.69
Uses novelty mean	1.38 ^v	1.35	1	2.12	0.02
Uses novelty max	2.77 ^v	3	1	4	0.08
Invention fluency	3.51 ^v	3	1	13	0.21
Invention novelty mean	2.51 ^v	2.52	1.25	3.78	0.05
Invention novelty max	3.17 ^v	3.33	1.67	4	0.07
Uses practicality	2.83 [^]	2.86	2.47	3	0.01
Invention practicality	3.43 [^]	3.50	1.5	4.33	0.04
RAT	0.56	0.55	0	1	0.02
Letters	0.50 ^v	0.52	0	0.78	0.01

^v $p < .05$ lower than Experiment 1; [^] $p < .05$ higher than Experiment 1

Alternative Uses

In contrast to Experiment 1, there was no effect of room size on mean fluency, $F(1,104)=0.4$, $p=.53$, with participants in the large rooms generating about the same number of uses ($M=20.3$, $SE=0.9$) as participants in the small room ($M=19.3$, $SE=1.2$), $d=0.13$ $[-0.28, 0.54]$. In contrast, similar to Experiment 1, mean trends for novelty of uses were in the hypothesized direction: however, the mean differences did not reach statistical significance. Mean novelty of uses was nonsignificantly higher in the large room ($M=1.4$, $SE=0.0$) compared to the small room ($M=1.3$, $SE=0.0$), $d=0.34$ $[-0.07, 0.75]$, $F(1,104)=2.8$, $p=.10$ (Fig. 8, left panel). Max novelty was marginally higher in the large rooms ($M=2.9$, $SE=0.1$) compared to the small room ($M=2.6$, $SE=0.1$), $d=0.37$ $[-0.03, 0.78]$, $F(1,104)=3.4$, $p=.07$ (Fig. 8, middle panel).

Similar to Experiment 1, there was an effect of room size on uses practicality, $F(1,104)=4.1$, $p=.04$, with participants in the larger rooms generating less practical uses ($M=2.8$, $SE=0.01$) than participants in the small room ($M=2.9$, $SE=0.02$), $d=-0.41$ $[-0.82, 0.00]$ (Fig 8, right panel).

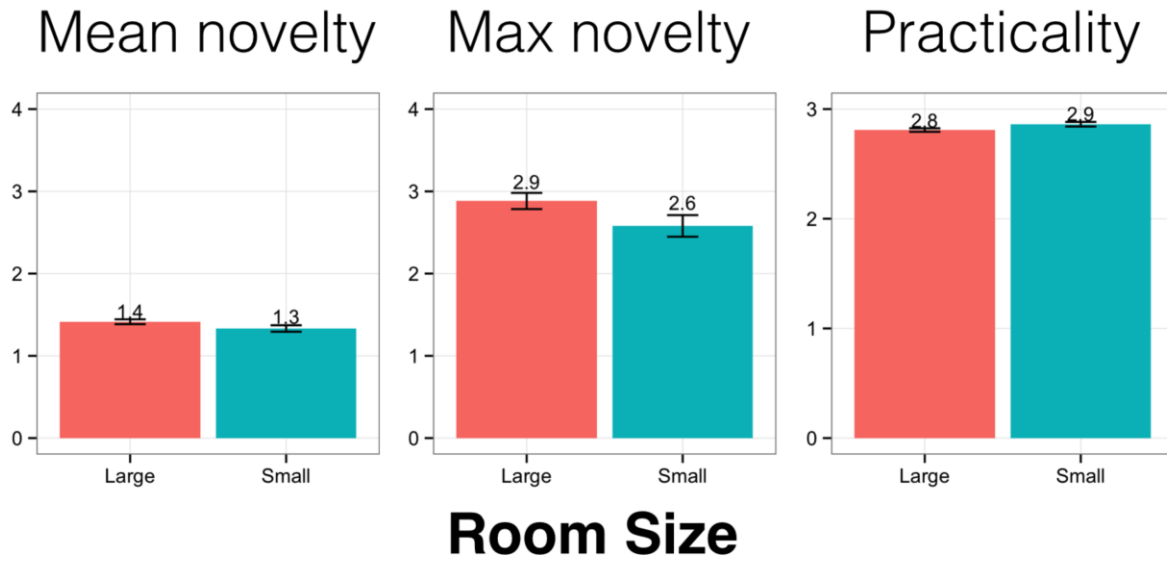


Figure 8. Novelty and practicality of alternative uses by room size in Experiment 2. Error bars are ± 1 SE.

Shape Invention

There was no effect of room size on mean fluency, $F(1,104)=0.9$, $p=.36$, with participants in the large room generating about the same number of inventions ($M=3.4$, $SE=0.3$) as participants in the small room ($M=3.8$, $SE=0.3$), $d=-0.19$ $[-0.59, 0.22]$. In contrast to Experiment 1, there were no reliable effects of room size on novelty of inventions: participants in the larger rooms had similar mean novelty scores ($M=2.5$, $SE=0.1$) as participants in the small room ($M=2.5$, $SE=0.1$), $d=0.14$ $[-0.26, 0.55]$, $F(1,104)=0.5$, $p=.49$. Similarly, participants in the larger rooms achieved about the same max novelty scores ($M=3.2$, $SE=0.1$) as participants in the small room ($M=3.1$, $SE=0.1$), $d=0.06$ $[-0.35, 0.46]$, $F(1,104)=0.1$, $p=0.79$.

There was no effect of room size on invention practicality, with participants in the large room generating inventions that were about as practical ($M=3.4$, $SE=0.1$) as participants in the small room ($M=3.5$, $SE=0.1$), $d=-0.15$ $[-0.56, 0.25]$, $F(1,104)=0.6$, $p=.46$.

RAT

There was no effect of room size, $F(1,104)=0.02$, $p=.90$. Participants in the larger rooms

had about the same mean proportion correct ($M=.56$, $SE=.03$) as participants in the small room ($M=.55$, $SE=.04$), $d=0.03$ $[-0.37, 0.43]$.

Letter Series

There was no effect of room size, $F(1,104)=0.07$, $p=.79$. Participants in the larger rooms had about the same mean proportion correct ($M=.50$, $SE=.02$) as participants in the small room ($M=.50$, $SE=.02$), $d=-0.05$ $[-0.46, 0.35]$.

Perceived Task Difficulty

In contrast to Experiment 1, there was no effect of room size on perceived difficulty for the uses task, $F(1,104)=0.0$, $p=.98$. Participants in the larger rooms self-reported the same levels of cognitive load ($M=5.1$, $SE=.18$) as participants in the small room ($M=5.1$, $SE=.24$), $d=-0.00$ $[-0.40, 0.41]$. Similarly, for the invention task, participants self-reported the same level of difficulty in the large ($M=5.9$, $SE=0.21$) and small rooms ($M=5.9$, $SE=0.28$), $d=-0.00$ $[-0.41, 0.40]$, $F(1,104)=0.0$, $p=0.98$.

Results were the same as Experiment 1 for the RAT and letters tasks. For the RAT, perceived difficulty was about the same in the large ($M=7.5$, $SE=0.15$) and small rooms ($M=7.3$, $SE=0.20$), $d=0.13$ $[-0.28, 0.53]$, $F(1,104)=0.4$, $p=0.53$. Similarly, perceived difficulty of the letters task was the same in the large ($M=0.50$, $SE=0.02$) and small rooms ($M=0.50$, $SE=0.02$), $d=-0.05$ $[-0.46, 0.35]$, $F(1,104)=0.1$, $p=0.79$.

Discussion

In Experiment 2, we sought to replicate and extend the findings from Experiment 1. See Figure 8 for a summary of the effects. We observed very similar patterns of effects for the alternative uses, RAT, and letter series tasks. Similar to Experiment 1, novelty (both mean and max) of uses was higher in the larger vs. small room, although the effect size was substantially

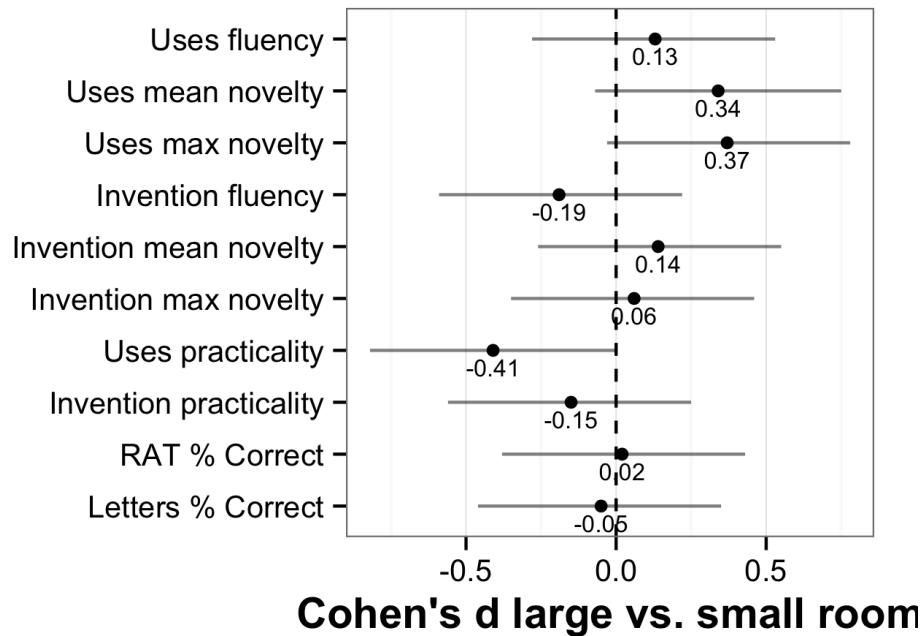


Figure 8. Summary of effects in Experiment 2. Error bars are 95% CIs.

smaller than Experiment 1 (approximately half the size). These trends are in the predicted direction of Hypothesis 1, and consistent with both the direct priming and concept activation accounts. Overall performance on the divergent measures were worse in both conditions in this experiment compared to Experiment 1, which may suggest floor effects. In the general discussion we further discuss possible reasons for the partial replication.

Partial support and replication was found for Hypothesis 2, with participants in the large room showing worse performance on the practicality measure of the alternative uses task compared to those in the small room. Also similar to Experiment 1, there was no effect of room size on the RAT or letter series task. However, perceived difficulty patterns did not replicate from Experiment 1. The failure to replicate the difference in perceived difficulty means that findings do not support the direct priming hypothesis more so than the concept activation account. In sum, we observed a partial replication of the results in Experiment 1 (mainly with novelty and practicality of uses).

GENERAL DISCUSSION

In this paper, we sought to explore the potential cognitive bases of the spatial metaphor that initial exploration of a creative space should be “broad”. Across two experiments, we tested the hypothesis that larger physical spaces facilitate divergent, but not convergent processes in problem-solving. Experiment 1 provided support for Hypothesis 1 and 2, and Experiment 2 partially replicated the findings. Smaller room sizes facilitated the generation of more practical uses of everyday objects across both experiments. Larger room sizes facilitated the generation of more novel uses of everyday objects (specifically mean and max novelty) in Experiment 1; although the same trends were seen in Experiment 2, the effects did not reach significance.

As noted in our discussion of the descriptive statistics in Experiment 2, there was a significant drop in performance across many of our measures from Experiment 1, suggesting that there may have been important differences in the two samples. One potential explanation is that Experiment 2 participants were exclusively undergraduate students participating for course credit, whereas Experiment 1 participants were paid volunteers and included a wider range of demographics (not just undergraduate students). Reduced motivation may have led to floor effects, potentially reducing the sensitivity of our measures. Motivation differences might have been especially important since we did not provide direct instructions to “be creative” in either experiment. For example, while the mean rated novelty of alternative uses was close to “not at all novel” across both experiments, the mean and variability was higher in Experiment 1 ($M=1.67$, $SD=0.42$) compared to Experiment 2 ($M=1.38$, $SD=0.24$). The generally low novelty scores (and high practicality scores) with relatively low variance suggest that participants were in general defaulting to more convergent processing. This observation is consistent with prior research on the “path of least resistance” in creative production (Ward, 1994), and other work that has shown

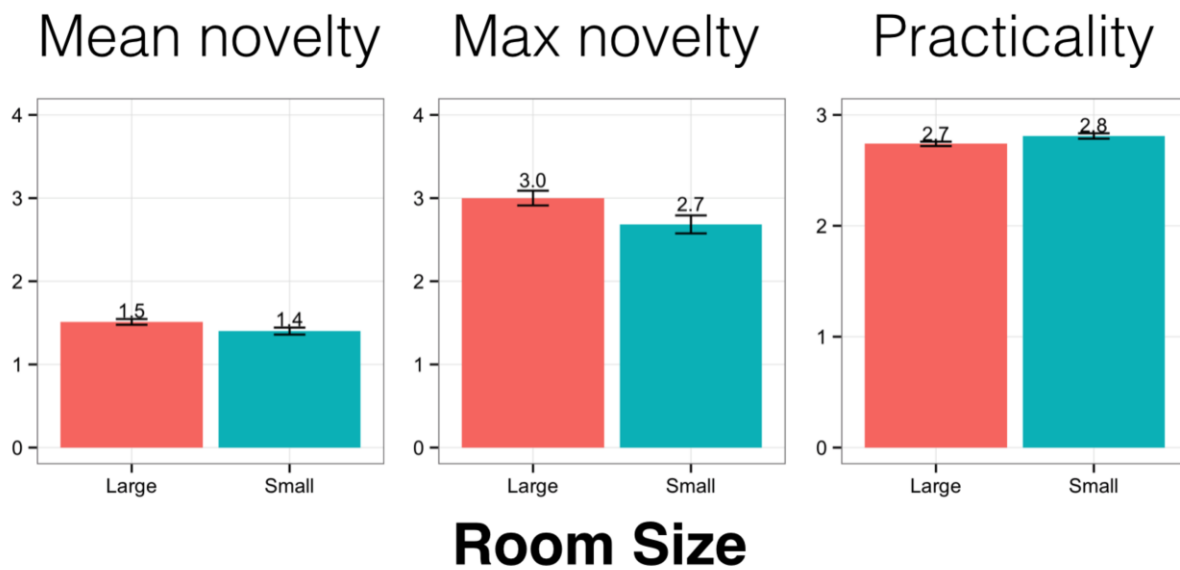


Figure 9. Novelty and practicality of alternative uses by room size, pooled across experiments. Error bars are ± 1 SE.

that instructions to “be creative” can yield substantial improvements to creative output (Nusbaum, Silvia, & Beaty, 2014): people generally need to expend cognitive effort to overcome initial biases towards less creative responses (e.g., using cognitive control to inhibit more accessible but less creative responses; Beaty, Silvia, Nusbaum, Jauk, & Benedek, 2014). In light of this, it may be useful to think of these results as describing the effects of room size on “default” problem solving (i.e., when participants are not necessarily actively trying to be creative in their responses).

We believe that pooling the data from the two experiments provides the clearest picture (e.g., robust across a wide range of participants and problem solving stimuli) of whether there is a relationship between room size and divergent and convergent problem solving processes. The pooled data indicate reliable evidence that larger physical spaces facilitate novelty and hinder practicality of solutions on the alternative uses task (see Figure 9). In the pooled data, mean novelty of uses is higher in the larger rooms ($M=1.5$, $SE=0.03$) compared to the small rooms ($M=1.4$, $SE=0.04$), $d=0.34$ [0.01, 0.67], $F(1,147)=4.13$, $p=.04$. Similarly, max novelty of uses is

higher in the larger rooms ($M=3.0$, $SE=0.09$) compared to the small rooms ($M=2.7$, $SE=0.11$), $d=0.37$ [0.04, 0.71], $F(1,147)=5.13$, $p=.03$. In contrast, mean practicality of uses was lower in the larger rooms ($M=2.7$, $SE=0.02$) compared to the small rooms ($M=2.8$, $SE=0.02$), $d=-0.38$ [-0.71, -0.04], $F(1,147)=5.11$, $p=.03$.

Our primary goal in this study is to document a psychological phenomenon: we provide an initial test of whether there is an association between room size and divergent problem solving performance. While this effect appears to be relatively small and may depend on having sufficiently motivated participants, some aspects of our results provide hints for future theoretical refinement. While our results from the problem solving measures (and non-replication of the hypothesized perceived task difficulty results from Experiment 1 to Experiment 2) are consistent with both a direct priming explanation (Hills et al. 2006, 2008, 2012), as well as the concept activation explanation (Meyers-Levy & Zhu, 2007), the affect results in Experiment 1 help to partially arbitrate between the explanations. The lack of effect on positive/negative affect is more consistent with a direct priming explanation, since concepts of “freedom” or “openness” are expected to engender more positive affect, whereas attunement to resource distribution patterns are not. We therefore suggest that, to the extent that this effect proves reliable, it may be a consequence of automatic attunement of semantic search patterns to search affordances in the physical environment. That is, people may be responding to the physical search affordances of the physical environment by defocusing their attention to enable broader search in semantic memory, shifting from a tight focus on a few highly relevant responses to considering more semantically distant and varied responses. However, alternative explanations are possible: for example, it is possible that the small room reminded people of traditional office environments, which may have invoked a “work schema” that primed more focused attention, impairing

divergent thinking. Further investigations are necessary to tease apart the psychological underpinnings of this effect.

Although we intended the letter series task and RAT as measures of convergent problem solving processes based on the nature of the response required (many = divergent, single = convergent), in hindsight, it is probably best to consider both the RAT and the letter series task as a more even mix of divergent and convergent processes than the fluency and novelty measures for the uses and invention tasks (which quite cleanly measure divergent processes). For example, in the RAT, one might first search broadly for possible meanings (strong and weak associates) of the target words and then only later converge on the common target that links across all three. Similarly, the letter series task may first require divergent search for several possible patterns before converging on the single, correct pattern to extrapolate. The mixture of both divergent and convergent processes in the letter series task and RAT might explain why we did not find a harmful effect of large rooms on letter series and RAT performance, and why we only found a harmful effect of large rooms on practicality of uses (which turned out empirically to be our cleanest measures of convergent processes, based on the intercorrelations between measures).

Our data have broader implications for the psychology of creative problem solving. For example, our observed strong negative correlations between novelty and practicality of uses corroborate prior arguments that originality and practicality in creative thought are cognitively at odds with each other (Goldenberg et al., 2013). Our results also have implications for how should we think about the RAT as a measure of creativity. In this study, we departed from a number of prior studies that have examined the divergent aspects of the RAT, for example studying the relationship between defocused attention and RAT performance (Aiello et al., 2012; Haarmann et al., 2012). However, the differing patterns of results and lack of correlation between

the RAT and divergent problem solving measures for the invention and uses task suggest that the RAT may involve more convergence than is typically described. Noting these findings might lead to more fruitful theoretical examinations of the relationship between the RAT and problem solving and creativity. For example, Goel and colleagues (2015) recently argued that insight problems (such as the RAT) are a subset of well-structured problems, while divergent problem solving tasks (such as the alternative uses and invention tasks) are a subset of ill-structured problems. Our results also echo a number of recent studies that have demonstrated the psychological separability of divergent and convergent problem solving processes (Chermahini & Hommel, 2010; Colzato et al., 2012; Hommel, Colzato, Fischer, & Christoffels, 2011; Oppezzo & Schwartz, 2014; Radel et al., 2015). We join these more nuanced arguments to call for more careful analyses of the components of creative performance (e.g., separating divergent vs. convergent processes) in future research on creativity.

REFERENCES

- Adánez, A. M. (2005). Does quantity generate quality? Testing the fundamental principle of brainstorming. *Spanish Journal of Psychology*, 8(2), 215–220.
<http://doi.org/10.1017/S1138741600005096>
- Aiello, D., Jarosz, A., Cushen, P., & Wiley, J. (2012). Firing the Executive: When an Analytic Approach to Problem Solving Helps and Hurts. *The Journal of Problem Solving*, 4(2).
<http://doi.org/10.7771/1932-6246.1128>
- Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology*, 45(2), 357.
- Antonenko, P., Paas, F., Grabner, R., & Gog, T. van. (2010). Using Electroencephalography to Measure Cognitive Load. *Educational Psychology Review*, 22(4), 425–438.
<http://doi.org/10.1007/s10648-010-9130-y>
- Beaty, R. E., Silvia, P. J., Nusbaum, E. C., Jauk, E., & Benedek, M. (2014). The roles of associative and executive processes in creative cognition. *Memory & Cognition*, 42(7), 1186–1197. <http://doi.org/10.3758/s13421-014-0428-8>
- Bhalla, M., & Proffitt, D. R. (1999). Visual-motor recalibration in geographical slant perception. *Journal of Experimental Psychology: Human Perception and Performance*, 25(4), 1076–1096.
- Bilalić, M., McLeod, P., & Gobet, F. (2008). Why good thoughts block better ones: the mechanism of the pernicious Einstellung (set) effect. *Cognition*, 108(3), 652–661.
- Bono, E. D. (1970). *Lateral Thinking: Creativity Step by Step*. New York: Harper Colophon.
- Bowers, K. S., Regehr, G., Balthazard, C., & Parker, K. (1990). Intuition in the context of discovery. *Cognitive Psychology*, 22(1), 72–110.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning.

Educational Researcher, 18(1), 32–42.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction.

Cognition and Instruction, 8(4), 293–332.

Chan, J., Fu, K., Schunn, C. D., Cagan, J., Wood, K. L., & Kotovsky, K. (2011). On the benefits and pitfalls of analogies for innovative design: Ideation performance based on analogical

distance, commonness, and modality of examples. *Journal of Mechanical Design*, 133,

081004. <http://doi.org/10.1115/1.4004396>

Chermahini, S. A., & Hommel, B. (2010). The (b)link between creativity and dopamine:

spontaneous eye blink rates predict and dissociate divergent and convergent thinking.

Cognition, 115(3), 458–65. <http://doi.org/10.1016/j.cognition.2010.03.007>

Colzato, L. S., Ozturk, A., & Hommel, B. (2012). Meditate to Create: The Impact of Focused-

Attention and Open-Monitoring Training on Convergent and Divergent Thinking.

Frontiers in Psychology, 3(April), 116.

Finke, R. A., Ward, T. B., & Smith, S. M. (1996). *Creative cognition: Theory, research, and applications*. Cambridge, MA: MIT Press.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American*

Psychologist, 52(1), 45–56.

Girotra, K., Terwiesch, C., & Ulrich, K. T. (2010). Idea generation and the quality of the best

idea. *Management Science*, 56, 591–605.

Goel, V., Eimontaite, I., Goel, A., & Schindler, I. (2015). Differential Modulation of

Performance in Insight and Divergent Thinking Tasks with tDCS. *The Journal of*

Problem Solving, 8(1). <http://doi.org/10.7771/1932-6246.1172>

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience* (Vol. ix).

Cambridge, MA, US: Harvard University Press.

Goldenberg, O., Larson, J. R., & Wiley, J. (2013). Goal Instructions, Response Format, and Idea Generation in Groups. *Small Group Research, 44*(3), 227–256.

<http://doi.org/10.1177/1046496413486701>

Greeno, J. G. (1994). Gibson's affordances. *Psychological Review, 101*(2), 336–342.

Greeno, J. G., & Middle-School Mathematics through Applications Project Group. (1998). The situativity of knowing, learning, and research. *American Psychologist, 53*(1), 5.

Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin, 53*(4), 267–293.

<http://doi.org/10.1037/h0040755>

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

Haarmann, H., George, T., Smaliy, A., & Dien, J. (2012). Remote Associates Test and Alpha Brain Waves. *The Journal of Problem Solving, 4*(2). <http://doi.org/10.7771/1932-6246.1126>

Hall, E. T. (1966). *The hidden dimension (1st ed.)* (Vol. xii). New York, NY, US: Doubleday & Co.

Hills, T. T. (2006). Animal foraging and the evolution of goal-directed cognition. *Cognitive Science, 30*(1), 3–41.

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal Foraging in Semantic Memory.

Psychological Review, 119(2), 431–440.

Hills, T. T., Todd, P. M., & Goldstone, R. L. (2008). Search in external and internal spaces: evidence for generalized cognitive search processes. *Psychological Science, 19*(8), 802–808.

- Hommel, B., Colzato, L. S., Fischer, R., & Christoffels, I. K. (2011). Bilingualism and creativity: benefits in convergent thinking come with losses in divergent thinking. *Frontiers in Psychology*, 2(November), 273. <http://doi.org/10.3389/fpsyg.2011.00273>
- Jang, J., & Schunn, C. D. (2012). Physical Design Tools Support and Hinder Innovative Engineering Design. *Journal of Mechanical Design*, 134(4), 041001.
- Kaplan, C., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology*, 22(3), 374–419.
- Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(6), 1534–1555.
- Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences*, 11(1), 16–22. <http://doi.org/10.1016/j.tics.2006.10.012>
- Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs*, 54(6), i–95. <http://doi.org/10.1037/h0093502>
- MacGregor, J., & Cunningham, J. (2009). The Effects of Number and Level of Restructuring in Insight Problem Solving. *The Journal of Problem Solving*, 2(2). <http://doi.org/10.7771/1932-6246.1062>
- Maclachlan, G. L., & Reid, I. (1994). *Framing and interpretation*. Melbourne, Australia: Melbourne University Press.
- Martindale, C. (1997). Creativity and connectionism. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The Creative Cognition Approach* (pp. 249–268). Cambridge, MA: MIT Press.
- Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review*, 69(3), 220–232.

- Mednick, S. A., & Mednick, M. T. (1967). *Examiner's manual: Remote Associates Test*. Boston, MA.
- Meyers-Levy, J., & Zhu, R. J. (2007). The Influence of Ceiling Height: The Effect of Priming on the Type of Processing That People Use. *Journal of Consumer Research*, *34*(2), 174–186.
- Moore, G. T., Lane, C. G., Hill, A. B., Cohen, U., & McGinty, T. (1994). *Recommendations for Child Care Centers. Revised Edition*. Center for Architecture and Urban Planning Research, University of Wisconsin-Milwaukee, Milwaukee, WI 53201-0413 (\$27.50 prepaid). Retrieved from <http://eric.ed.gov/?id=ED417797>
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ.
- Nokes-Malach, T. J., & Mestre, J. P. (2013). Toward a Model of Transfer as Sense-Making. *Educational Psychologist*, *48*(3), 184–207. <http://doi.org/10.1080/00461520.2013.807556>
- Nokes, T. J. (2009). Mechanisms of knowledge transfer. *Thinking & Reasoning*, *15*(1), 1–36.
- Nusbaum, E. C., Silvia, P. J., & Beaty, R. E. (2014). Ready, set, create: What instructing people to “be creative” reveals about the meaning and mechanisms of divergent thinking. *Psychology of Aesthetics, Creativity, and the Arts*, *8*(4), 423–432. <http://doi.org/10.1037/a0036549>
- Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. In M. T. Keane & K. J. Gilhooly (Eds.), *Advances in the psychology of thinking* (Vol. 1, pp. 1–44). Hertfordshire, UK: Harvester Wheatsheaf.
- Oppezzo, M., & Schwartz, D. L. (2014). Give your ideas some legs: The positive effect of walking on creative thinking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 1142–1152. <http://doi.org/10.1037/a0036577>

- Parnes, S. J., & Meadow, A. (1959). Effects of “brainstorming” instructions on creative problem solving by trained and untrained subjects. *Journal of Educational Psychology, 50*(4), 171–176.
- Radel, R., Davranche, K., Fournier, M., & Dietrich, A. (2015). The role of (dis) inhibition in creativity: Decreased inhibition improves idea generation. *Cognition, 134*, 110–120.
- Sawyer, R. K. (2012). *Explaining creativity: the science of human innovation* (2nd ed.). New York: Oxford University Press.
- Scherr, R. E., & Hammer, D. (2009). Student Behavior and Epistemological Framing: Examples from Collaborative Active-Learning Activities in Physics. *Cognition and Instruction, 27*(2), 147–174. <http://doi.org/10.1080/07370000902797379>
- Shah, J. J., Millsap, R. E., Woodward, J., & Smith, S. M. (2012). Applied Tests of Design SkillsPart 1: Divergent Thinking. *Journal of Mechanical Design, 134*(2), 021005–021005–10. <http://doi.org/10.1115/1.4005594>
- Simon, H. A., & Kotovsky, K. (1963). Human acquisition of concepts for sequential patterns. *Psychological Review, 70*(6), 534–546.
- Simonton, D. K. (2011). Creativity and discovery as blind variation: Campbell’s (1960) BVSR model after the half-century mark. *Review of General Psychology, 15*(2).
- Terwiesch, C., & Ulrich, K. T. (2009). *Innovation tournaments: Creating and selecting exceptional opportunities*. Boston, MA: Harvard Business Press.
- Torrance, E. P. (1988). The nature of creativity as manifest in its testing. In R. J. Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives* (pp. 43–75). New York, NY: Cambridge University Press.
- Wallas, G. (1926). *The art of thought*. New York, NY.

- Ward, T. B. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology*, 27(1), 1–40.
- Ward, T. B. (1998). Analogical distance and purpose in creative thought: Mental leaps versus mental hops. In K. J. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in Analogy Research: Integration of Theory and Data from the Cognitive, Computational, and Neural Sciences* (pp. 221–230). Sofia, Bulgaria.
- Warr, A., & O'Neill, E. (2005). Understanding Design As a Social Creative Process. In *Proceedings of the 5th Conference on Creativity & Cognition* (pp. 118–127). New York, NY, USA: ACM. <http://doi.org/10.1145/1056224.1056242>
- Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, 45(4), 1191–1207. <http://doi.org/10.3758/s13428-012-0314-x>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–70.
- Wiley, J. (1998). Expertise as mental set: The effects of domain knowledge in creative problem solving. *Memory & Cognition*, 26(4), 716–730.
- Wiley, J., & Jarosz, A. F. (2012). Working Memory Capacity, Attentional Focus, and Problem Solving. *Current Directions in Psychological Science*, 21(4), 258–262. <http://doi.org/10.1177/0963721412447622>

APPENDIX A: FULL LIST OF ITEMS FOR RAT TASK

1. Broken-Clear-Eye
2. Skunk-Kings-Boiled
3. Coin-Quick-Spoon
4. Manners-Round-Tennis
5. Barrel-Root-Belly
6. Notch-Flight-Spin
7. Cloth-Sad-Out
8. Foot-Collection-Out
9. Stick-Light-Birthday
10. Sore-Shoulder-Sweat
11. Blank-White-Lines
12. Bass-Complex-Sleep
13. Gold-Stool-Tender
14. Playing-Credit-Report
15. Time-Hair-Stretch
16. Cracker-Union-Rabbit
17. Off-Trumpet-Atomic
18. Rabbit-Cloud-House
19. Room-Blood-Salts
20. Salt-Deep-Foam
21. Square-Cardboard-Open
22. Water-Tobacco-Stove
23. Ache-Hunter-Cabbage
24. Chamber-Staff-Box
25. High-Book-Sour
26. Lick-Sprinkle-Mines
27. Pure-Blue-Fall
28. Square-Telephone-Club
29. Shopping-Washer-Picture
30. Note-Dive-Chair
31. Magic-Plush-Floor
32. Inch-Deal-Peg

APPENDIX B: FULL LIST OF ITEMS FOR LETTER SERIES TASK

1. aaabbbccdd _____
2. atbataatbat _____
3. abmcdmefmghm _____
4. defgefghfghi _____
5. qxapxbqxa _____
6. aduacuaeubuafua _____
7. mabmbcmcdm _____
8. urtustuttu _____
9. abyabxabwab _____
10. rscdstdetuef _____
11. npaoqapraqsa _____
12. wxaxybyzczadab _____
13. jkqrklrslmst _____
14. pononmnmllk _____
15. lmzmlymnx _____
16. efsferfgq _____
17. cdqdcpdeo _____
18. ijwjivjku _____