Novel Conceptual Combination in Individuals and Groups

Michael DiBiase

Hofstra University

Abstract

This experiment examines the effects of group size, levels of relatedness, and levels of typicality on creativity and abstractness of labels and exemplars in a novel creativity task. The effects of relatedness and typicality on creativity have been studied before (Mobley, Doares, & Mumford, 1992), but never in conjunction with group size and abstractness. - Seventy-two undergraduate students were randomly assigned to work either in groups of three or individually. Each group or individual was handed the same eight sets of twelve words. The relatedness and typicality of these groups of words has already been predetermined by Mobley et al. Participants provided a category label that best fit the words presented to them, with a one-sentence description of the new category, and were instructed to generate as many exemplars for the new category as they could. The twelve stimulus words come from three different categories, although the categories were not explicitly labeled for the participants. Relatedness refers to the relationship between the three categories of items. Typicality refers to the relationship of each exemplar to its own category. Mobley has cited evidence in previous studies that atypicality and a low relatedness lead to more creative responses. The prediction was that groups would be more creative, especially when participants generated new exemplars that fit into the category the participants had created. My results find that for individuals, related categories with typical exemplars produced the most cohesive exemplars. In conjunction with Mobley's findings, my results indicate that for groups, related categories yielded exemplars that were rated as the most cohesive regardless of typicality. My experiment

shows that three participants working as a group generated a larger number of unique exemplars than an individual set of three participants.

Novel Conceptual Combination in Individuals and Groups

The effect of conceptual combination on creativity in individuals has been widely studied in various applications. However, little research has been conducted on groups performing conceptual combination tasks. By understanding how category-relatedness, exemplar-typicality, and group dynamics interact with regards to creativity, many applications in the business world become available. Managers would choose very carefully the way in which a memo would be written to spark the most creativity from his or her employees. A better understanding of the mechanics of creativity may give insight into whether brainstorming from a group will derive more original, quality, and abstract/concrete answers than would an individual.

The term conceptual combination refers to the way in which one unifies words, thoughts, or ideas into something novel. If someone is presenting a business proposal, they will draw upon their previous knowledge and experience, and combine it with the needs of their boss or client to hopefully come up with a successful idea; this is conceptual combination at work. Individual creativity can result from prior knowledge and situations that one experiences; conceptual combination can be used to merge this knowledge to come up with a novel solution to a problem. One of the most relevant studies, and the one on which this was based on was conducted by Mobley, Doares, and Mumford (1992). They stress the importance that prior knowledge has in the availability of schemata (categorical structures), which can be called upon to form a novel or creative idea. These categories, fruits for example, are made up of exemplars, such as grapes, oranges, or apples. Mobley et al. (1992) describe creative thought as a combination and

reorganization of these categories. In their experiment, they tested the existence and operation of the combination and reorganization process in creative problem solving, and the interaction of the combination and reorganization process with other processes such as encoding (the transferring of data into memory) and construction (the process of building a plan or idea from prior knowledge) and their effect on quality and originality. Quality was defined as categories, exemplars, or stories that were appropriate with respect to the context provided by the taxonomic or given categories used in the problem, and yielded a coherent or meaningful organization of relevant elements (Mobley et al. 1992). For example, a response of quality corresponding with given exemplars such as train, plane, and car would be labeled "transportation." On the other hand, a response lacking quality for the given exemplars (train, plane, and car) would be the label "animals" because it is not relevant with respect to the context of the presented words. Originality was defined as a novel response which went beyond the manifest stimulus context, and provided a vehicle for organizing presented material in a new way (Mobley et al. 1992). The label "transportation" for the given exemplars (train, plane, and car) lacks originality because it is not novel. However, a response with originality to these given words might be the label "things that cause motion sickness."

To test these concepts, Mumford et al. presented participants with twelve words, and they were to generate a category label that encompassed the words. Participants then had to give a one-sentence description of the category. Next, participants were to list as many exemplars as possible that fit under that new category. After that, participants had

to write a one-paragraph story on the label they created for three out of the fifteen problems they were presented. The twelve words were varied in three separate ways: one group of four words would be either biological or artifactual, one group would be typical or atypical, and one group would be related or unrelated. The term biological refers to categories that are biological objects (e.g., spices, trees, fish). The term artifactual refers to categories that are inanimate objects (e.g., medical tools, sports equipment, linens). Typicality and atypicality refer to the degree to which given exemplars are typical of a certain category. For example, typical words for the category "liquids" are: beer, water, coffee, and milk. Atypical words for the category "liquids" are: beer, paint, blood, and milk. Relatedness and unrelatedness refer to the degree to which the three categories presented exemplars are similar. Related categories are: things having to do with a car (seat, tire, brakes, wheel), things having to do with sports (glove, baseball, baseball hat, football), and things having to do with exercise (bicycling, running, swimming, and lifting weights). Unrelated categories are: things having to do with a bed (pillow, mattress, headboard, linen), things having to do with water activities (boating, sailing, scuba diving, swimming), and things having to do with surgery (scalpel, needle, clamp, rubber gloves) (Mobley et al. 1992).

Mumford et al, found that the participants' category labels and exemplars received the highest creativity scores when presented with unrelated as opposed to related categories. (Mobley et al. 1992). In the comparison

between typicality and atypicality, the results suggest that more quality responses are given when typical exemplars are presented. The presence of unambiguous information facilitates the construction of coherent, meaningful categories (Mobley et al. 1992). The comparison between the biological and artifact categories yielded interesting results. Biological categories, being explicit and well-defined, resulted in higher scores of category quality. Artifact categories, being of human origin or production, allowed greater flexibility and resulted in higher scores of category originality.

Another study further investigated how combination and reorganization play an integral role in the production of novel ideas (Mumford et al. 1997). Combination and reorganization refer to the way in which people combine and reorganize existing knowledge structures or conceptual categories. The investigators build upon the ideas of combination and reorganization as a means of understanding a problem situation, and creating new ideas (Mumford et al. 1997). The experiment turned into a question of how people react to similar and diverse problem situations. The hypothesis was that people will use feature mapping when confronted with similar categories that have overlapping features, and that when asked to work with diverse categories, people will use metaphors to guide the combination and reorganization process (Mumford et al. 1997). Feature mapping refers to a process involving four basic operations: (a) the central properties of the key features of category members are identified (e.g. birds fly), (b) the features or properties identified for one category will be mapped onto the features of the other categories (e.g. birds fly and have feathers, whereas airplanes fly and are made of metal), (c) shared features will be used to construct a new category, and (d) additional features or

properties of category members will then be identified through elaboration (Mumford et al. 1997). The use of metaphors, turning to broader analogical implication of these features rather than solely the concrete meaning, may be one way in which people combine and reorganize information. For example, birds fly, but flight may be held to represent freedom or escape (Mumford et al. 1997). The two questions the investigators sought to answer were: do effective combination and reorganization efforts contribute to performance on an independent set of creative problem-solving tasks drawn from different domains, and is performance on combination and reorganization tasks related to people's ability to solve complex, novel problems (Mumford et al. 1997).

Two experiments were run to investigate these questions: the first was a category-exemplar generation problem followed by a series of creative problem-solving tasks. In the first task, participants were to create an advertising campaign for a product called the 3-D Holographic Television. They were given a description of the product and were told to write a two or three paragraph description of what they deemed a successful marketing survey, a television advertisement, and a magazine advertisement (Mumford et al. 1997). As in the last study, the participants' responses were rated for quality and originality utilizing a detailed rubric for appropriating scores to responses. For the originality scale, the rubric deemed a score of 1 appropriate for a "Vague description of ad: 'Upbeat, youthful theme.' Nondescriptive placement" (Mumford et al. 1997). The description for a response deserving the score of 5 was much different. "Ad tells the 3D TV story using 'flat/Columbus' theme with modern twist…(e.g., Columbus in lounge, with a Walkman, reading the Titanic, Crew enters and yells, 'Its true, the earth is flat.') Placement

discussed." (Mumford et al. 1997). A similar rubric was provided for scoring the quality of responses.

The second type of problem-solving task involved management and public policy problems. Participants were presented with a description of a problem, and they were to write a one or two paragraph description of how they would solve the issue. Their responses were then scored for quality and originality using similar rubrics as the previous problem-solving task.

In order to determine how metaphors contribute to the creation of more original products when subjects are presented with diverse stimulus categories, two between-subjects manipulations were made in the conditions of task performance. The first manipulation took place in the category-exemplar experiment. The lists of words participants received were either scored as having high-relatedness or low-relatedness (Mumford et al. 1997). The second manipulation was intended to influence the strategies participants used while they completed the category-exemplar problems. In the first condition, participants were instructed to identify features of each category and then map these features onto each other. The second condition had the same instructions, however, participants were told to think about what the features may represent before starting work on the actual category-exemplar problems (Mumford et al. 1997). These instructions allowed for the broader analogical interpretations of features to surface, utilizing the metaphorical contribution.

The results appeared to support the notion that people's ability to combine and reorganize categories to generate new concepts is related to the ability to produce

original, high-quality products on creative problem-solving tasks (Mumford et al. 1997). Participants who successfully did this tended to produce higher scores of quality and originality. The significance is that combination and encoding, the transformation of information to be stored in the memory, might be one of the most important influences on creative problem solving. Concerning my experiment, comparing individual responses to group responses, one might assume that a group of three participants would have a larger collective knowledge base than an individual participant, giving the group more information to draw upon.

Mobley et al. (1992) and Mumford et al. (1997) developed some of the most relevant theories pertaining to this experiment. The theory of the effect of atypicality on creative problem solving is based on previous research documenting that typical exemplars are processed more accurately and rapidly than atypical ones. As a result, the experimenters theorized that information encoding and category search may be affected by varying typicality. They believe that a more extensive conscious search process, stimulated by those atypical exemplars, would contribute to problem solving by creating a wider bank of knowledge for a person to draw upon (Mobley et al. 1992).

The next theory Mumford et al. put forth was based on previous research regarding the diversity of knowledge structures, or the way in which prior experiences and information are categorized. Relatedness was varied to manipulate the combination and reorganization process. Previous research has shown that the generation of novel problem solution is facilitated by the application of diverse knowledge structures which are not commonly associated or linked to one another (Mobley et al. 1992). This is

what led the investigators to hypothesize that unrelated exemplars would encourage original and quality responses from participants.

Several individuals working together to create new labels and exemplars would not only have their own knowledge bases to draw upon, but also the knowledge bases of the group. Furthermore, an idea from one individual might serve as a catalyst for another individual's thought process.

However, those individuals having trouble merging unrelated given exemplars might abstract their label responses in order to encompass all twelve words. The level of abstractness refers to how vague or broad response labels and exemplars become in order to successfully include the twelve given exemplars. One might turn to an abstract label such as "things in the world" to encompass unrelated words.

Kuhn (1970) found that anomalous or inconsistent observations often served as a stimulus for problem construction. More centrally, however, the proposed solution to the problem was often associated with a young person's reorganization of the concepts used in past attempts to understand related problems (Mobley et al. 1992). Furthermore, Rothenberg (1986) and his colleagues provided experimental evidence which speaks to the potential significance of reorganization. This evidence suggests that the linkage of previously unrelated concepts often provide a basis for later creative achievement. Rothenberg (1986), Rothenberg & Sobel (1980), and Sobel & Rothenberg (1980) defined *homospatial thinking* as deliberately and consciously superimposing two or more images on each other to create a new form (Mobley et al. 1992). Unrelatedness facilitates the same superimposing to create a novel solution for the combination of unlike exemplars.

Mobley et al. (1991) delve further into variables that influence a creative outcome. The investigators believe that the systematic combination and reorganization of existing categories in memory greatly affect novel solutions one might use to solve a problem. Ward et al. (2001) stress the cognitive processes and structures as being more important than creativity itself. These processes attempt to explain creativity and where it comes from. One topic in particular described by Ward et al. has significant relevance to the experiment at hand. Conceptual combination is that the merging of two or more concepts can result in a novel entity that is more than the simple sum of its component parts (Ward et al. 2001). Ward et al. use the tangible example of hydrogen and oxygen merging to form water, a result that neither parent element contained alone. The category-exemplar experiment is a perfect example of putting this theory to the test. The label that is created to encompass the exemplars would represent the resulting combination of a successful merge of two or more concepts. The new exemplars that fit into the category label would represent the emergence of new properties that were not present before the combination.

Ward et al. (1995) describe the phenomenon of *structured imagination*. When asked to be as imaginative as possible while creating animals from another planet, college students tended to produce creatures highly similar to earth animals in that they were bilaterally symmetric, and possessed ordinary appendages and sense organs (Ward et al. 1995). This study underscores the reutilization of familiar images and ideas when problem solving. It is the combining of these pre-existing notions that has the potential to

create a new entity. Groups as opposed to individuals might allow for a larger selection of prior information, resulting in a higher probability of the creation of something novel.

Nemeth et al. (2003) further outline the advantages and disadvantages that groups possess in a creativity task. They cite the desire of consensus as a serious pitfall for groups. This desire leads to premature closure, such as that evidenced by research on groupthink (Nemeth et al. 2003). The tendency towards consensus can lead those involved in brainstorming activities to bypass their own opinions to follow those of the majority. This is due to the common notion that the majority must be correct. Despite these disadvantages, groups do have a secret weapon in the form of dissent. It is a liberator of thought and, perhaps more important, a stimulus to divergent and creative thought (Nemeth et al. 2003). It only takes one dissenter in a group to overturn the power of the majority. The dissenter stimulates thinking about an issue from multiple angles; something an individual is lacking. By enabling other members of the group to view something from a number of angles, the dissenter has allowed the group to take advantage of divergent thought, resulting in more novel solutions. Stasser & Birchmeier (2003) introduce the concept of the hidden profile, information that can divert a group from choosing the common idea over the unique. Overpowering the consensus built upon shared prior knowledge, a group can analyze a hidden profile, constructed from divergent prior knowledge, to create an innovative and plausibly superior outcome. The concepts of groupthink, dissenters, and the hidden profile play a large role in my experiment because the combination of knowledge-bases of group participants can be either suppressed or stimulated, affecting the quality and originality of group responses.

Nemeth et al's concepts of groupthink and the dissenter may be reflected in the success or failure of my participating groups. Ideally, groups have the potential to achieve higher scores than individuals due to the advantage of a wider knowledge-base. However, the common problem of groupthink may hinder this natural advantage even with the presence of a dissenter or hidden profiles.

Milliken & Martins (1996) describe group diversity as being dual-natured. Diversity, thus, appears to be a double-edged sword, increasing the opportunity for creativity as well as the likelihood that group members will be dissatisfied and fail to identify with the group (Milliken & Martins, 1996). This reinforces the idea that the composition of the group can affect the outcome of its responses. The effectiveness of a group's creative process is two-fold: divergent thinking, the generation of different ideas and possible options, and convergent thinking, the process of selecting the most suitable solution for a collective goal. If this method is utilized, participants working in groups may have an advantage in producing a category-label for the presented exemplars in my experiment. With the ability to create divergent options but also select the best convergent solution, groups may have a potential advantage over individuals.

The experiment I conducted is built around Mumford et al's category-exemplar problem experiments. I have built upon the original theory of the effect of atypicality that states that typical exemplars are processed more accurately and rapidly than atypical ones, and as a result, information encoding and category search may be affected by varying typicality (Mobley et al. 1992) (Mumford et al. 1997). A more extensive

conscious search process, stimulated by those atypical exemplars, would contribute to problem solving by creating a wider bank of knowledge for a person to draw upon. I have adopted these theories, and applied them to the comparative responses of groups and individuals. I used Mumford et al's category/exemplar setup while altering the scoring variables to include cohesiveness, fitness, abstractness, and relatedness. Individuals as well as groups would be presented with twelve words which have been pre-determined to be related/unrelated and typical/ atypical. Participants were then to come up with a one-sentence label that encompassed the twelve words, and then as many new words that would fit under the new label they just created. Raters then scored each label based on how well it fit the twelve words, its abstractness, and its originality. Exemplars were scored based on how well they fit the label, how well the exemplars seemed to have come from one category as opposed to multiple categories, level of abstractness, and originality. These scores are the dependent variables. The independent variables are whether the exemplars are related or unrelated, whether the exemplars are biological or artifactual, whether the exemplars are typical or atypical, and whether an individual or group is performing the task.

My prediction is that with regard to the labels, groups will score higher on how well the label fits the twelve original words, groups will show a lower abstractness level, and will show more originality than individuals. I also believe that unrelated categories and atypical exemplars will show these results more often than related categories and

typical exemplars. I base my predictions on previous research which has shown time after time that creative problem-solving comes from previous knowledge that is then reconstructed. Unrelated categories have been shown to spark a wider search of previous knowledge which may be indicative of increased creativity. Although Mumford et al's research did not show a significant relationship between atypicality and creativity, I believe when comparing individuals to groups, there will be a notable difference. For example, atypical words should lead to more creative labels and exemplars for groups. It is logical that a group of three people will have a broader spectrum of prior knowledge to draw upon than an individual working by himself/herself. Furthermore, I predict that individuals will have more abstract responses because he/she is drawing from a single knowledge base; whereas, a group will have three knowledge bases offering a larger amount of concrete experiences and information to draw upon.

My prediction with regard to the exemplars is that groups will score higher with how well the exemplars fit the label, how well the exemplars seem to come from one category as opposed to three separate ones, groups will show a lower abstractness level, and higher originality. I predict that unrelated categories and atypical exemplars will show these results as opposed to related categories and typical exemplars. I believe the abstractness scores will be greater in individuals because if an individual is unable to come up with a label to describe the presented words, they would most likely come up with a very broad label. I believe groups will be able to come up with a more targeted label due to the larger knowledge base they are drawing from, and in-turn will have more targeted and accurate exemplars. The goal of my experiment is to reaffirm the role of

conceptual combination and reorganization, and, furthermore, display the potential of greater creativity in groups as opposed to individuals.

Method

Participants

Seventy-two Hofstra University Psychology 001 students volunteered via an online sign-up program to participate for course credit. The first thirty-six participants were placed into groups of three, forming twelve groups. Only one group was tested at a time. The other thirty-six participants were tested as individuals. Two individuals were tested at the same time.

Materials

Participants were all shown the same eight sets of twelve words, but the order in which the sets were presented was randomly assigned. Each group was presented with a different order of word sets, resulting in 12 distinct patterns. As a means for comparison, three individuals were assigned to the same order that one group used. So although each group of three participants had a unique order of the eight sets of words, every three individuals had the same order.

Design

This study was a 2 x 2 x 2 x 2 mixed factorial design. The independent variables were participation in a group or as an individual, relatedness of the categories, typicality of the exemplars presented, and whether the presented exemplars were biological or artifactual. The dependent variables were ratings of the originality, quality, and abstractness of the new category description and exemplars, and the number of

exemplars generated. The originality, quality, and abstractness were rated independently by three clinical psychology graduate students of Hofstra University. The participants' responses were rated on a scale of 1-5. The instructions for the raters are reproduced in Appendix A.

Procedure

The procedure was similar to the one used by Mobley et al. (1992) with the exception that half of the participants worked in groups and the abstractness of the responses was measured. Once everyone had arrived, each participant was handed an informed consent form (Appendix B) which was then read aloud. The instructions (Appendix C) were then handed out and read aloud. The participants were instructed that they would be presented with eight sets of 12 words. The eight sets of exemplars are shown in Appendix D. Each set would be presented on a separate piece of paper. For each set, the participants were to come up with a one-sentence label or description that best encompassed the 12 words. They were then instructed to write down as many new words that fit into the new category they had just created. The groups had four minutes to work on each set of words, and were warned when only one minute remained. After the instructions were read, the participants had a chance to ask any questions they had. The experimenter read a simple example problem to make sure that the participants were all confident with the experiment. The sample exemplars were apple, pear, and banana. The participants were then shown two possible labels (fruits, or fruits you eat for a snack). They were then shown possible new exemplars (strawberry, peach, and orange).

The group was then asked to nominate a recorder. This person had the responsibility of recording the group's one-sentence label, as well as all of the new words.

The 36 participants who worked as individuals came in sets of two to the sessions. They signed a consent form and were instructed they would be working by themselves. The two individuals were instructed to sit on opposite ends of the table so they would not interfere with each other. Aside from that, they followed the same instructions and procedures as those working in groups.

After the experiment was over, the participants were debriefed. They were informed that the study was designed to see how different words affect creativity. Participants were informed that various combinations of related and unrelated categories, as well as typical and atypical exemplars were used to try and evoke or suppress original, quality, and abstract responses. They were told about the manipulation of groups versus individuals to see the effects of these independent variables on creativity. Participants were than asked if they would like to receive the results of the study by e-mail.

Results

An inter-rater reliability test was used to confirm that the three raters scored the participants' responses with consistency. Their scores were to reflect the definitions of fitness, cohesiveness, abstractness, originality we agreed upon in a previous meeting. Rater 1 ended up being the only results I could use in my analysis. Rater 2 and 3 primarily used the extremes of the 1-5 scale while Rater 1 felt comfortable giving a range of scores. It is clear that Rater 2 and 3 were not comfortable distinguishing between the

middle-range scores, and thus I was forced to discard their data. The results below are, as a result, strictly from the scores I received from Rater 1.

The mean fitness ratings for the category labels are given in Table 1. The category labels for biological exemplars were rated as having greater fitness (3.67) than the labels for artifactual categories (3.37), $F(1,22)=7.92$, p=.01. The fitness of the labels was rated significantly higher for related exemplars (3.69) than for those that were unrelated (3.35), $F(1,22)=11.54$, $p=.003$. Fitness ratings did not differ significantly between groups (3.54) and individuals (3.5), $F(1,22)=.068$, $p=.797$. Fitness ratings for typical exemplars (3.58) did not differ significantly from ratings for atypical exemplars (3.46), $F(1,22)=1.42$, $p=.246$. There was a significant two-way interaction (Figure 1) between relatedness and type of category (biological/artifactual), $F(1,22)=21.32$, $p=.000$. Biological, typical exemplars resulted in a lower fitness scores than artifactual, typical exemplars. However, biological, atypical exemplars resulted in higher scores of fitness than artifactual, atypical exemplars.

The mean abstractness ratings for the category labels are given in Table 2. Abstractness ratings did not differ significantly between groups (3.25) and individuals (2.93), $F(1,22)=3.6$,$p=.071$. There was a significant two-way interaction (Figure 2) between typicality and type of category (biological/artifactual), $F(1,22)=7.63$, $p=.011$ Typicality did not have much of an effect on the abstractness ratings of biological exemplars. However, artifactual exemplars' ratings for abstractness increased from typical to atypical. There was also a significant three-way interaction between type of category (biological/artifactual), relatedness, and individuals (Figure 3) vs. groups

(Figure 4), $F(1,22)=5.38$, $p=.03$, such that in both groups and individuals, the rating for a biological exemplar dropped from related to unrelated categories. Although individuals started with much lower scores in abstractness for artifactual exemplars in related categories, their abstractness surpassed that of biological exemplars in unrelated categories.

The mean originality ratings for the category labels are given in Table 3. The originality of the labels was rated significantly higher for atypical exemplars (2.67) than for those that were typical (2.25), $F(1,22)=14.51$, $p=.001$. The originality of the labels was rated significantly higher for unrelated exemplars (2.75) than for those that were related (2.16), $F(1,22)=15.88$, $p=.001$. Originality ratings did not differ significantly between groups (2.41) and individuals (2.51), $F(1,22)=.312$, $p=.582$. There was a significant two-way interaction (Figure 5) between type of category (biological/artifactual) and relatedness, $F(1,22)=12.63$, $p=.002$. Biological exemplars in related categories showed slightly higher originality than artifactual exemplars. Biological and artifactual exemplars in unrelated categories both showed increased originality scores; however artifactual exemplars increased by a much greater quantity than biological exemplars.

The mean fitness ratings for the category exemplars are given in Table 4. The fitness of the exemplars was rated significantly higher for typical exemplars (4.04) than for those that were atypical (3.74), $F(1,22)=9.52$, $p=.005$. The fitness of the exemplars was rated significantly higher for biological exemplars (3.99) than for those that were artifactual (3.79), $F(1,22)=5.2$, $p=.033$. The fitness of the exemplars was rated

significantly higher for related categories (4.1) than for those that were unrelated (3.68), $F(1,22)=16.15$, $p=.001$. Fitness ratings did not differ significantly between groups (3.98) and individuals (3.8), $F(1,22)=2.12$, $p=.159$.

The mean cohesiveness ratings for the category exemplars are given in Table 5. The cohesiveness of the exemplars was rated significantly higher for related categories (3.67) than for those that were unrelated (3.06), $F(1,22)=53.77$, $p<.000$. Cohesiveness ratings did not differ significantly between groups (3.38) and individuals (3.36), $F(1,22)=.019$, $p=.891$. There was a significant three-way interaction between typicality, relatedness, and individuals (Figure 6) and groups (Figure 7), $F(1,22)=5.711$, $p=.026$, such that there is an increase in cohesiveness in related categories with atypical exemplars, and a decrease in cohesiveness for unrelated categories with atypical exemplars for groups. In individuals, there is a decrease in cohesiveness for related categories with atypical exemplars and an increase in cohesiveness for unrelated categories with atypical exemplars.

The mean abstractness ratings for the category exemplars are given in Table 6. Abstractness ratings did not differ significantly between groups (2.48) and individuals (2.49), $F(1,22)=.009$, $p=.923$. There was a significant two-way interaction (Figure 8) between type of category (biological/artifactual), and relatedness, $F(1,22)=6.74$, $p=.016$, such that there is an inverse relationship between category type and relatedness. Artifactual exemplars in related categories resulted in a lower abstractness score than biological exemplars in related categories. However, artifactual exemplars in unrelated

categories resulted in higher scores of abstractness than biological exemplars in unrelated categories.

The mean originality ratings for the category exemplars are given in Table 7. The originality of the exemplars was rated significantly higher for related categories (3.26) than for those that were unrelated (3.05), $F(1,22)=8.00$, $p=.01$. The originality of the exemplars was rated significantly higher for groups (3.84) than for individuals (2.46), $F(1,22)=32.15$, $p=.000$.

The mean number of exemplars as a function of a unique label is given in Table 8. The number of exemplars per novel label was significantly higher for groups (23.26) than for individuals (10.16), $F(1,22)= p<.05$. The mean number of total exemplars label is given in Table 9.

There was a significant two-way interaction (Figure 9) between type of category (biological/artifactual) and relatedness, $F(1,22)=16.86$, $p=.000$. There was a significant three-way interaction between typicality, type of category (biological/artifactual), and relatedness, $F(1,22)=6.73$, $p=.017$.

Six correlation tests were run that compared the exemplars to labels for fitness, abstractness and originality. This was done for both groups and individuals in an attempt to explain why certain scores were lower than others.

When the fitness exemplar scores were compared to the fitness label scores for the individuals, a moderately strong (r =.42, p=.000) correlation was found. The comparison between abstractness scores for individuals resulted in a weak correlation (r

=.24, p=.01). When the originality exemplar scores were compared to the originality label scores for the individuals, no correlation (r = -.05, p=.32) was found.

For groups, a weak correlation (r =.22, p=.03) was found between exemplar fitness and label fitness. There was no correlation (r =.007. p=.47) found between exemplar abstractness and label abstractness. There was also a very weak correlation (r =.14, p=.101) between exemplar originality and label originality.

Discussion

Although the results failed to meet all aspects of the hypothesis, groups and individuals did show a sizable difference in the originality of their exemplars. Individuals were fairly consistent in their scores from labels to exemplars, however, groups were not. Whereas the originality of groups' labels were scored as relatively poor, their exemplars received scores twice as original as the labels. Because the scores for individuals and groups were pretty similar for the labels, the advantage of a group seems to shine when producing exemplars. This may be attributed to poor group dynamics such as the need to consent, the belief that the majority is correct, and simply going with the first response that comes to mind; all of these things would attribute to a poor label. Another possibility is that someone within the group may be afraid to suggest something strange. This potentially changes with the production of exemplars. Whereas all the factors mentioned before nullify the biggest advantage a group has over an individual (a larger knowledge structure to use in conceptual combination), exemplars may wipe that slate clean. Once a label is created, whether original or not, participants in a group can start drawing on their own knowledge.

A possible reason that individuals showed consistent originality ratings from label to exemplar may be due to the fact that they boxed themselves into a label they were unable to expand upon. This might result in mundane and unoriginal exemplars. If a label was significant to each group member in a different way, the potential for an original group of exemplars has risen. This is further exemplified in the quantity of exemplars produced per novel category label. Groups produced more than twice the amount of novel exemplars than individuals did. This again may be a result of the greater knowledge structure groups had to draw upon. However, it should be noted that if the total number of exemplars are considered, regardless of label repetition, individuals did produce slightly more than groups. Previous studies have produced the opposite results. The common finding is that individuals will produce exemplars of higher quality while groups create exemplars in higher quantity. I would be interested to see if I find the same results on a larger-scale.

One idea is that the overlap of knowledge ordinarily seen in three people may be reduced in a situation like this. If a particular label has a completely different significance to each person in the group, the knowledge base may have expanded overall. A possible way to test this would be to find out what individuals within a group were thinking about when coming up with new exemplars. A way to test this would be to add another section to the experiment where individuals explain how they came up with their responses. By gaining insight into the size of the knowledge structure a participant is using, it's possible to prove a group will have a larger knowledge-base to draw upon overall.

The interaction between typicality, relatedness, and groups vs. individuals for cohesiveness ratings exemplifies an inverse relationship which appeared a number of times throughout the results. For individuals, related categories with typical exemplars produced the most cohesive exemplars. Unrelated categories with typical exemplars resulted in the least cohesive scores. This may be due to the fact that individuals had a hard time combining the unrelated categories, resulting in exemplars that seemed to come from three individual categories as opposed to one. For groups, related categories yielded exemplars that were rated as the most cohesive regardless of typicality, while atypical exemplars from unrelated categories were rated as the lowest. This may indicate another advantage groups have over individuals; when it comes to the cohesiveness of exemplars, groups seem unaffected by typicality when given related exemplars. Relatedness seems to play a larger role in cohesiveness possibly because it acts as a larger hindrance. Both groups and individuals found the typicality manipulation of the category more challenging than the relatedness manipulation. This observation suggests that participants had difficulty with concepts they were less familiar with as opposed to concepts that varied.

Another interaction that displayed how a variable failed to produce the same effect across every level of groups is the interaction between category type (biological/artifactual), relatedness, and groups vs. individuals. Category type turned out to be an important factor in many interactions. In both groups and individuals, biological exemplars were rated with higher abstractness scores for related categories as opposed to unrelated. When it came to artifactual exemplars, individuals' scores started low in

related categories and escalated in unrelated categories. When presented with an unrelated category, many individuals simply described each one of the three categories in one label (animals, liquids, and trees). This resulted in a very abstract, yet unoriginal response. The groups' scores for artifactual exemplars remained unchanged in related and unrelated categories. My theory is that groups did not have to resort to the technique of simply providing a label that stated the three categories. The groups' expanded wider knowledge-base may have provided them a wider spectrum to combine unrelated categories instead of simply resorting to the individuals' method.

As I mentioned before, category type (biological/artifactual) played a much larger role than I had expected. The first interaction I encountered was between biological/artifactual and relatedness. Artifactual exemplars exhibited a much lower fitness rating for related categories than biological exemplars. For unrelated categories, however, where artifactual exemplars remained relatively constant, biological exemplars dropped drastically. Fitness ratings for biological exemplars may have dropped due to a greater difficulty to combine unrelated categories compared to artifactual. To me, there are more daily applications for inanimate objects than biological ones. This may lend the advantage to a wider knowledge structure for artifactual as opposed to biological exemplars.

Although the combination of relatedness and biological exemplars resulted in low scores, category type and relatedness with respect to originality is a different story. Both biological and artifactual exemplars had low originality ratings for labels in related categories. For unrelated categories, artifactual exemplars scored extremely well. I

believe that unrelated categories present the opportunity for creativity much more than atypical exemplars. In this study, relatedness was varied to manipulate the combination and reorganization process. Previous research has shown that the generation of novel problem solutions was facilitated by the application of diverse knowledge structures which are not commonly associated or linked to one another (Mobley et al. 1992). This especially rings true for artifactual exemplars. I found it much easier to come up with a label for the artifact exemplars than the biological ones because there are so many more scenarios that are applicable with inanimate objects than animate ones.

Within Mobley et al's 1992 experiment, the variable of quality was used. In my own experiment, fitness, the measure of the appropriateness of the label to the original exemplars, replaces this variable. I will be comparing quality and fitness due to their similarity in meaning. As expected for Mobley et al. 1992 the relatedness manipulation produced significant main effects for the measures of category quality and exemplar quality. Their related condition yielded higher quality for both exemplars (related (4.06) vs. unrelated (3.94)) and categories (related (3.45) vs. unrelated (3.27)). My results were similar, showing fitness of the labels as significantly higher for related exemplars (3.69) than for those that were unrelated (3.35). Mumford et al's results showed higher ratings of originality for unrelated exemplars in both categories (unrelated (2.98) vs. related (2.60)) and exemplars (unrelated (2.92) vs. related (2.81)). Although my originality ratings for unrelated categories (unrelated (2.75) vs. related (2.16)) paralleled Mumford et al's findings, my originality ratings for unrelated exemplars were not in agreement. My originality ratings for the related condition (3.26) were significantly greater than the

unrelated condition (3.05). Another way in which my findings differ from Mumford et al's is in the response to typicality/atypicality. Higher quality categories were obtained in high typicality condition (typical (3.47) vs. atypical (3.25)) (Mobley et al, 1992). Within my experiment, the fitness ratings of categories for typical exemplars (3.58) did not differ significantly from ratings for atypical exemplars (3.46). Another point of comparison is biological vs. artifactual. Yielding corresponding results, Mobley et al's 1992 experiment and my own indicate that given biological exemplars result in higher ratings of quality/fitness in exemplars and category labels. A final difference between Mumford et al's study and my own was the effects of biological/artifactual on originality. Where they found a discernible response showing artifact categories producing higher originality for categories and exemplars, my own results were unclear.

Although some trends were apparent in my results, I would have liked to have been able to use more than one of my raters' scores. Despite meeting with my raters to discuss how one would score abstractness, originality, cohesiveness, and fitness, only one of them really showed an understanding after looking at their scores. A rubric system, a standardized document outlining the appropriate scores that should be given to varying responses, as in Mumford et al. 1997 may have made it easier for my raters to show more consistency in their scores. A rubric would have also made it easier for my raters to use the entirety of the scale, instead of just the extremes. It is evident that it was difficult for them to differentiate between the middle scores. From the results I was able to use, it is clear that fitness, abstractness, cohesiveness, and originality are not as cut and dry as my hypothesis had predicted.

The relationships between typicality, category type (biological/artifactual), relatedness, and individuals/groups are complicated and overlapping. The interactions I found do show a strong tendency for groups to be rated as having more original exemplars, and a much larger number of exemplars per novel label. That was the most exciting result as many previous studies have found individuals to be higher generators of exemplars than groups. I believe the trend towards group superiority that my results have shown can be attributed to the wider knowledge-base groups possessed over individuals.

References

Baughman, W. A., Costanza, D. P., Maher, M. A., Mumford, M. D., & Supinski, E. P. (1997). Process-Based Measures of Creative Problem-Solving Skills: Category Combination. *Creativity Research Journal,* 10 (1), 59-71.

Mobley, M. I., Doares, L. M., & Mumford, M. D. (1992). Process Analytic Models of Creative Capacities: Evidence for the Combination and Reorganization Process. *Creativity Research Journal,* 5 (2), 125-155.

Doares, L. M., Mobley, M. I., Mumford, M. D., Reiter-Palmon, R., & Uhlman, C. E. (1991). Process Analytic Models of Creative Capacities. *Creativity Research Journal*, 4 (2), 91-122.

Kuhn T. S. (1970) The Structure of Scientific Revolutions. Chicago: Chicago University Press.

Milliken, F. J., & Martins, L. (1996). Searching for common threads: Understanding the multiple effects of diversity in organizational groups. *Academy of Management Review, 21*, 402-433.

Nemeth, J. C., & Nemeth-Brown, B. (2003). Better than Individuals? The Potential Benefits of Dissent and Diversity for Group Creativity. In P. B. Paulus and B. A. Nijstad (Eds.), *Group creativity: Innovation through collaboration* (pp. 137-159). New York: Oxford University Press.

Rothenberg, A. (1986). Artistic creation as stimulated by superimposed versus combined-composite visual images. *Journal of Personality and Social Psychology, 50*, 370-381.

Rothenberg, A., & Sobel. R. S. (1980). Creation of literary metaphors as stimulated by superimposed versus separated visual images. *Journal of Mental Imagery, 4*, 37-91.

Smith, S. M., Vaid, J., & Ward, T. B. (2001). Conceptual Structures and Processes in Creative Thought. *Creative Thought: An Investigation of Conceptual Structures and Processes*, 1-27.

Sobel, R. S., & Rothenberg, A. (1980). Artistic creation as stimulated by superimposed versus separated images. *Journal of Personality and Social Psychology, 39*, 953-961.

Stasser, G., and Birchmeier, Z. (2003). Group creativity and collective choice. In P. Paulus and B. Nijstad (Eds.), *Group Creativity* (pp. 132-172), Oxford University Press.

Ward, T. B., Finke, R. A., & Smith, S. M. (1995). *Creativity and the mind: Discovering the genius within*. New York: Plenum Press.

Appendix A

Instructions for Raters

You will be given two sets of 72 participant response sheets. One set will contain participant labels for a list of 12 words coming from 3 categories; the other set will contain participants list of new exemplars for the category labels.

First, rate the labels.

Flip through a random subset of response sheets to get an idea of how well the labels fit the words listed at the top of the page, the level of abstractness of the labels, and the level of originality of the labels. Your ratings will be made using a five-point scale. Write your ratings on the back of each response sheet.

For each response sheet, first rate the fit and abstractness of the label.

(1) How well does the label describe the set of twelve words?

1 2 3 4 5
(not at all) (excellent fit)

(2) How abstract is the label?

1 2 3 4 5
(concrete) (abstract)

When you have rated the fitness and abstractness of each label, then go through the labels again (in a new random order) and rate the originality of each label (the degree to which the label is novel and provides a new or unexpected way of organizing the words).

(3) How original is the label?

1 2 3 4 5
(not at all original) (very original)

Next, rate the exemplars.

Flip through a random subset of response sheets to get an idea of how well the exemplars fit the label written at the top of the page, the degree of cohesiveness, the level of abstractness of the exemplars, and the level of originality of the exemplars. Your ratings

will be made using a five-point scale. Write your ratings on the back of each response sheet.
For each response sheet, first rate the fit and abstractness of the exemplar.

(1) How well do the exemplars describe the label?

1 2 3 4 5
(not at all) (excellent fit)

(2) How cohesive are the exemplars? That is to what degree do they seem to have come from a single category as opposed to multiple categories?

1 2 3 4 5
(not at all cohesive/multiple categories) (quite cohesive/ single category)

(3) How abstract are the exemplars?

1 2 3 4 5
(rigid) (flexible)

When you have rated the fitness, cohesiveness, and abstractness of the exemplars, then go through the exemplars again (in a new random order) and rate the originality of the exemplars (the degree to which it is novel).

(4) How original are the exemplars?

1 2 3 4 5
(not at all original) (very original)

Appendix B

Informed Consent Form

This experiment examines people's reactions to different combinations of words. During this experiment, you will be asked to play a word game, according to the directions in the attached packet. The experiment should last approximately 45 minutes.

Your results will remain anonymous and will be used strictly for research. No identifying information will be kept. Your performance on this task will have no effect on your course grade.

Your participation in this experiment is voluntary and you may stop participating at any time and for any reason. If you choose to stop participating, you will not be penalized in any way.

The purpose and hypotheses of the experiment will be explained when you have completed this experimental session. You will be able to ask any questions you have regarding the purpose of the experiment at this time.

Your signature on this page indicates that you have read the information on this page and agree to participate in the experiment. Let the experimenter know if you have any questions at this time.

Signature_____ Date_____

Researcher information:
Student researcher: Mr. Michael DiBiase, 516-578-7501,
Faculty Sponsor: Dr. Vincent Brown

Appendix C

Instructions:

In this experiment you will be presented with eight sets of twelve words. Each set is presented on a separate page. For each set, you are to come up with a brief label or description that best describes the category representing the entire set of words. After you come up with the category label or description, you are to come up with as many examples as you can that fit under the new category. You will have four minutes to work on each set of words. The experimenter will warn you when you have one minute remaining.

Here is a simple example of the procedure using just three words (the actual task will be a little more difficult):

words: Apple Pear Banana

label/description: "fruit" or "fruits you eat for a snack"

other examples fitting the category: strawberry, peach, orange, etc.

Please let the experimenter know if you have any questions before proceeding with the experiment.

Appendix D

Unrelated, Typical, Biological

Fox	Beer	Aspen
Zebra	Paint	Maple
Tiger	Blood	Elm
Panther	Milk	Birch

Use the space provided to write your one sentence label for the 12 words above.

Use the space provided to come up with as many new words to fit into the label you created above.

Unrelated, Typical, Artifact

Pillow	Boating	Scalpel
Mattress	Sailing	Needle
Headboard	Scuba Diving	Clamp
Linen	Swimming	Rubber Gloves

Use the space provided to write your one sentence label for the 12 words above.

Use the space provided to come up with as many new words to fit into the label you created above.

Related, Atypical, Biological

Nicotine	Water	Champagne
Benzedrine	Blood	Brandy
Heroin	Paint	Sherry
Cocaine	Beer	Vodka

Use the space provided to write your one sentence label for the 12 words above.

Use the space provided to come up with as many new words to fit into the label you created above.

Unrelated, Atypical, Artifact

Footboard	Boating	Gas Mask
Canopy	Ballet	Needle
Headboard	Ice Skating	Tweezers
Linen	Swimming	Rubber Gloves

Use the space provided to write your one sentence label for the 12 words above.

Use the space provided to come up with as many new words to fit into the label you created above.

Related, Typical, Artifact

Seat	Glove	Bicycling
Tire	Baseball	Running
Brakes	Baseball Bat	Swimming
Wheel	Football	Lifting Weights

Use the space provided to write your one sentence label for the 12 words above.

Use the space provided to come up with as many new words to fit into the label you created above.

Related, Typical, Biological

Codeine	Water	Whiskey
Penicillin	Coffee	Brandy
Heroin	Milk	Rum
Cocaine	Beer	Vodka

Use the space provided to write your one sentence label for the 12 words above.

Use the space provided to come up with as many new words to fit into the label you created above.

Related, Atypical, Artifact

Light	Glove	Horseback Riding
Horn	Mask	Running
Brakes	Hockey Stick	Jumping
Wheel	Football	Lifting Weights

Use the space provided to write your one sentence label for the 12 words above.

Use the space provided to come up with as many new words to fit into the label you created above.

Unrelated, Typical, Biological

Lion	Beer	Pine
Rhinoceros	Water	Maple
Tiger	Coffee	Oak
Panther	Milk	Birch

Use the space provided to write your one sentence label for the 12 words above.

Use the space provided to come up with as many new words to fit into the label you created above.

Table 1

Label Fitness as a Function of Typicality, Relatedness, Category Type (Biological/Artifactual), and Group/Individual

Biological	Typical	Atypical	Typical	Atypical
	Individuals		Groups	
Related	4.19	4.00	4.17	3.83
Unrelated	3.61	3.06	3.08	3.42
	Individuals		Groups	
Artifactual				
Related	3.22	3.28	3.42	3.42
Unrelated	3.36	3.25	3.58	3.42

Table 2

Label Abstractness as a Function of Typicality, Relatedness, Category Type (Biological/Artifactual), and Group/Individual

Biological	Typical	Atypical	Typical	Atypical
	Individuals		Groups	
Related	3.06	2.94	3.50	3.25
Unrelated	2.69	2.92	3.08	3.50
	Individuals		Groups	
Artifactual				
Related	2.69	2.69	3.25	3.42
Unrelated	2.72	3.28	2.83	3.12

Table 3

Label Originality as a Function of Typicality, Relatedness, Category Type (Biological/Artifactual), and Group/Individual

Biological	Typical	Atypical	Typical	Atypical
	Individuals		Groups	
Related	2.44	2.36	1.92	2.08
Unrelated	2.44	2.75	1.92	2.83
	Individuals		Groups	
Artifactual				
Related	1.94	2.28	1.75	2.5
Unrelated	2.64	3.19	2.92	3.33

Table 4

Exemplar Fitness as a Function of Typicality, Relatedness, Category Type (Biological/Artifactual), and Group/Individual

Biological	Typical	Atypical	Typical	Atypical
	Individuals		Groups	
Related	4.11	4.03	4.33	4.25
Unrelated	3.64	3.61	4.00	3.67
	Individuals		Groups	
Artifactual				
Related	3.81	3.61	4.25	4.00
Unrelated	3.97	3.47	3.92	3.42

Table 5

Exemplar Cohesiveness as a Function of Typicality, Relatedness, Category Type (Biological/Artifactual), and Group/Individual

Biological	Typical	Atypical	Typical	Atypical
	Individuals		Groups	
Related	3.83	3.67	3.83	3.92
Unrelated	2.78	3.19	3.42	2.92
	Individuals		Groups	
Artifactual				
Related	3.56	3.42	3.42	3.42
Unrelated	3.08	2.97	3.08	3.00

Table 6

Exemplar Abstractness as a Function of Typicality, Relatedness, Category Type (Biological/Artifactual), and Group/Individual

Biological	Typical	Atypical	Typical	Atypical
	Individuals		Groups	
Related	2.31	2.58	2.58	2.92
Unrelated	2.50	2.53	2.33	2.42
	Individuals		Groups	
Artifactual				
Related	2.28	2.67	2.25	2.00
Unrelated	2.69	2.81	2.75	2.50

Table 7

Exemplar Originality as a Function of Typicality, Relatedness, Category Type (Biological/Artifactual), and Group/Individual

Biological	Typical	Atypical	Typical	Atypical
	Individuals		Groups	
Related	2.44	2.94	3.92	4.00
Unrelated	2.33	2.36	3.92	3.58
	Individuals		Groups	
Artifactual				
Related	2.67	2.56	4.12	4.00
Unrelated	2.81	2.78	3.83	3.33

Table 8

Total Number of Exemplars as a Function of Unique Category Labels, Typicality, Relatedness, Category Type (Biological/Artifactual), and Group/Individual

Biological	Typical	Atypical	Typical	Atypical
	Individuals		Groups	
Related	22.25	24.92	18.25	23.41
Unrelated	31.92	27.75	25.67	24.75
	Individuals		Groups	
Artifactual				
Related	33.00	26.00	26.92	24.75
Unrelated	28.25	26	21.92	20.42

Table 9

Total Number of Exemplars as a Function of Typicality, Relatedness, Category Type (Biological/Artifactual), and Group/Individual

Biological		Typical	Atypical		Typical	Atypical
	Individuals			Groups		
Related		25.33	27.75		17.92	22.5
Unrelated		35.58	29.5		25.42	24.08
	Individuals			Groups		
Artifactual						
Related		36.17	28.33		26.83	24.83
Unrelated		30.5	27.5		22	20.25

Figure Caption

Figure 1. Two-way interaction between relatedness and category type (biological/artifactual) for the fitness of labels.

Figure 2. Two-way interaction between typicality and category type (biological/artifactual) for the abstractness of labels.

Figure 3. The top graph is a three-way interaction between relatedness, category type (biological/artifactual) and individuals for the abstractness of labels. The bottom graph is a three-way interaction between relatedness, category type (biological/artifactual) and groups for the abstractness of labels.

Figure 4. Two-way interaction between relatedness and category-type (biological/artifactual) for the originality of labels.

Figure 5. The top graph is a three-way interaction between typicality, relatedness, and individuals for the cohesiveness of exemplars. The bottom graph is a three-way interaction between typicality, relatedness, and groups for the cohesiveness of exemplars.

Figure 6. Two-way interaction between category-type (biological/artifactual) and relatedness for the abstractness of exemplars.

Figure 7. Two-way interaction between category-type (biological/artifactual) and relatedness for the number of new exemplars.

Figure 1.

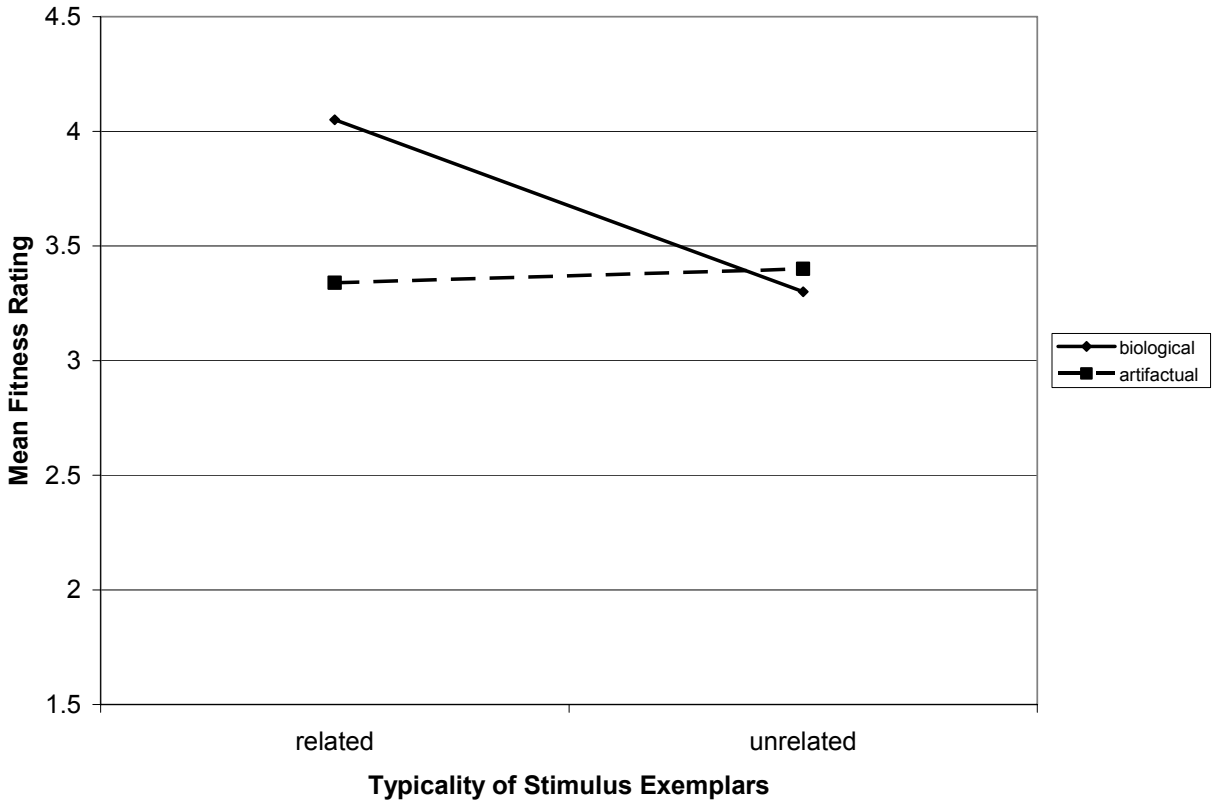

4.5
4
3.5
3
2.5
2
1.5
Mean Fitness Rating
related
unrelated
Typicality of Stimulus Exemplars
biological
artifactual

Figure 2.

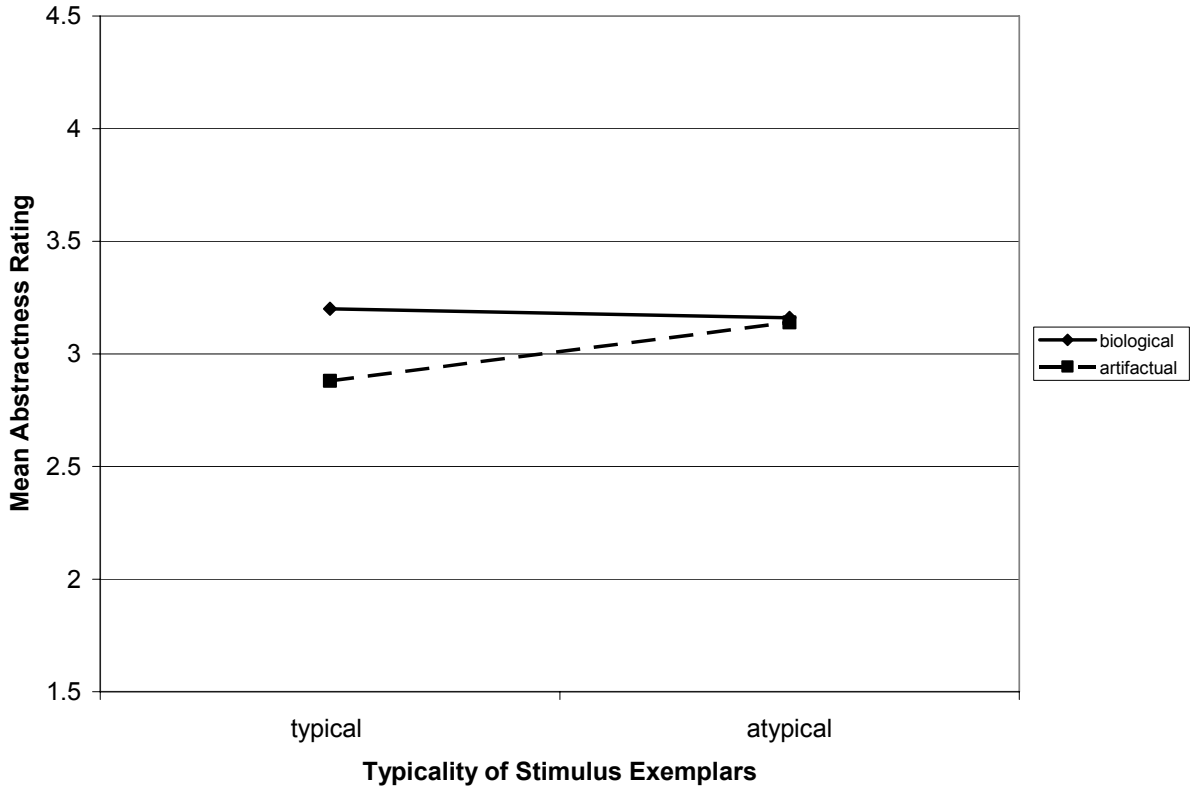
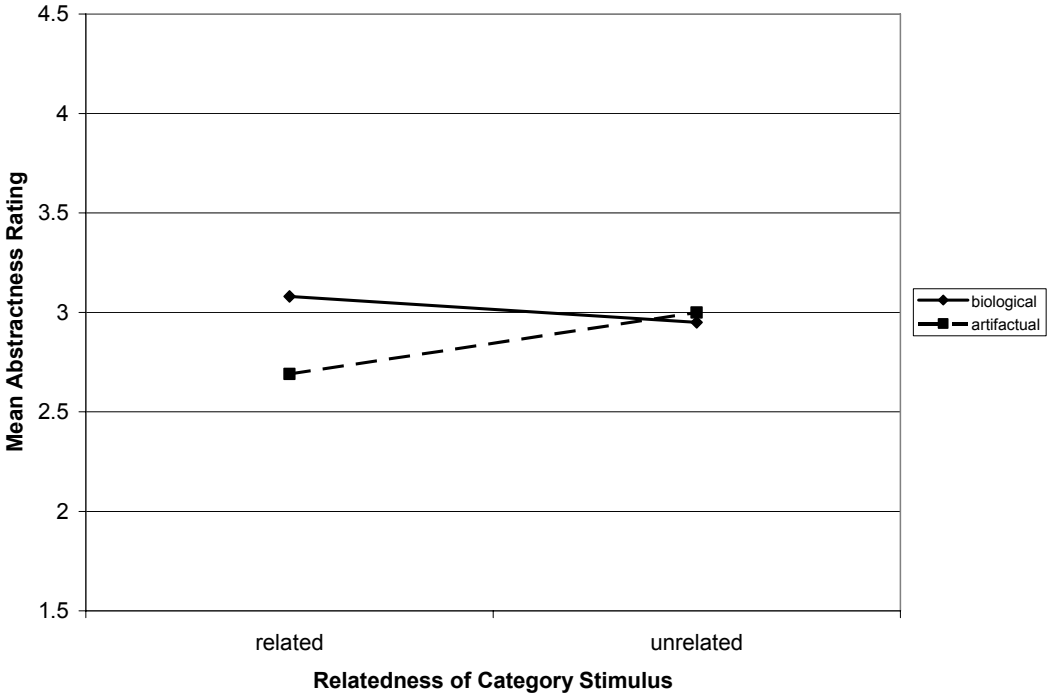

4.5
4
3.5
3
2.5
2
1.5
Mean Abstractness Rating
typical
atypical
Typicality of Stimulus Exemplars
biological
artifactual

Figure 3.


4.5
4
3.5
3
2.5
2
1.5
Mean Abstractness Rating
related
unrelated
Relatedness of Category Stimulus
biological
artifactual

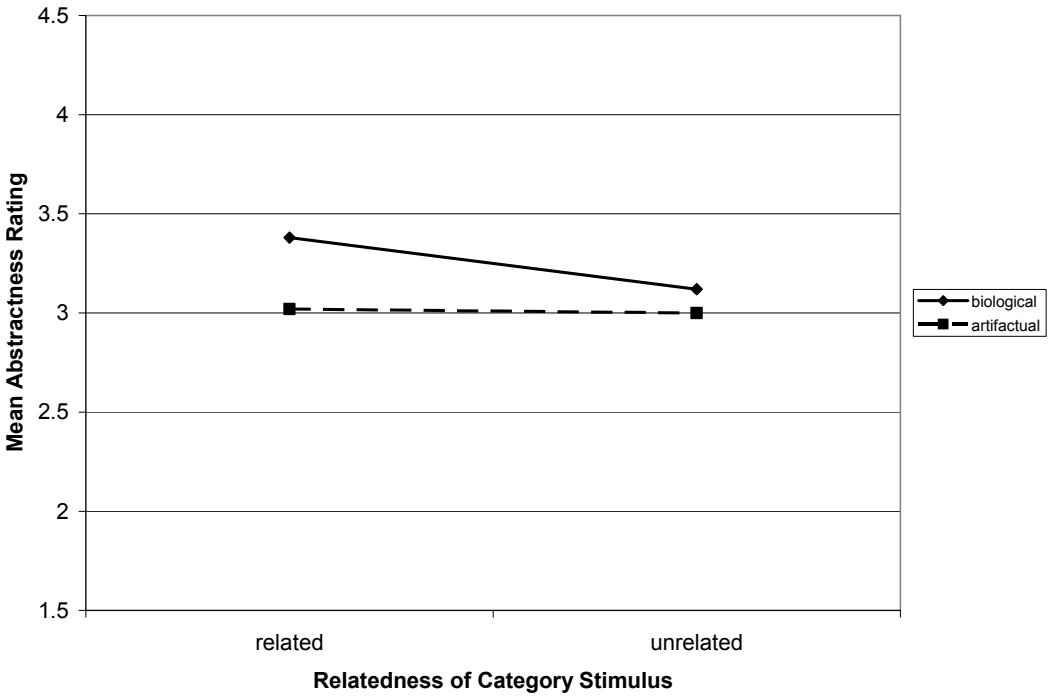

4.5
4
3.5
3
2.5
2
1.5
Mean Abstractness Rating
related
unrelated
Relatedness of Category Stimulus
biological
artifactual

Figure 4.

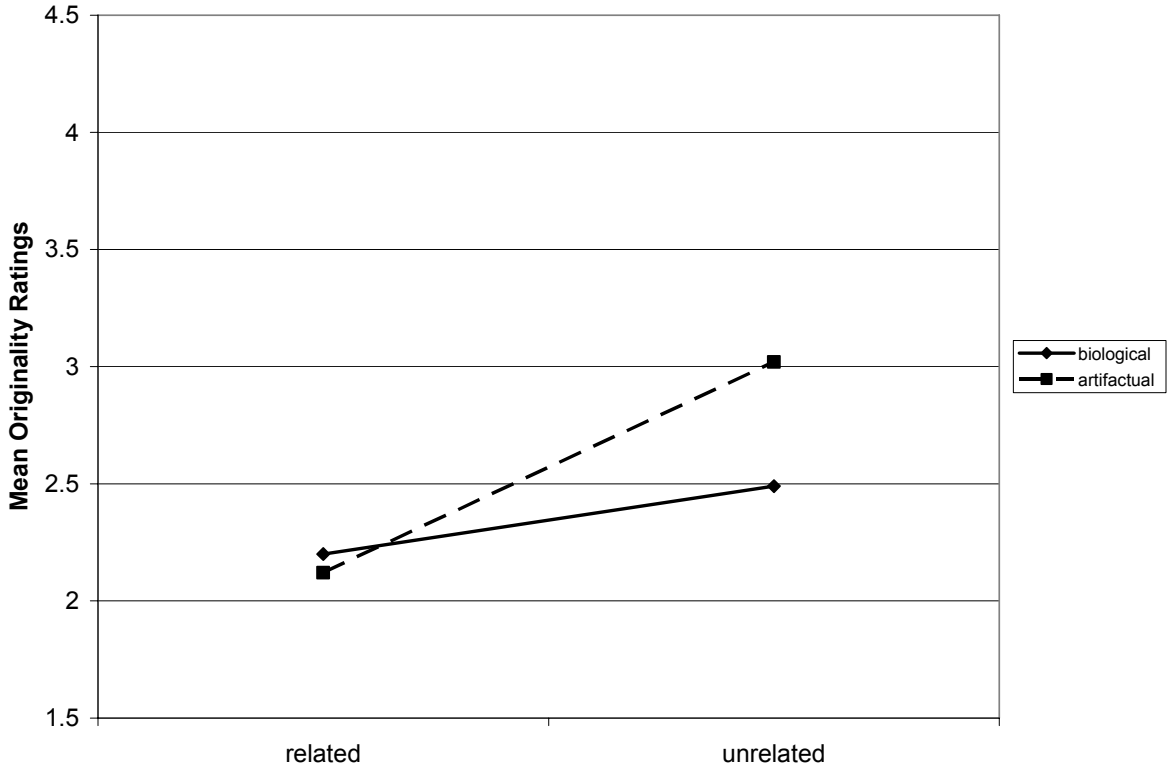
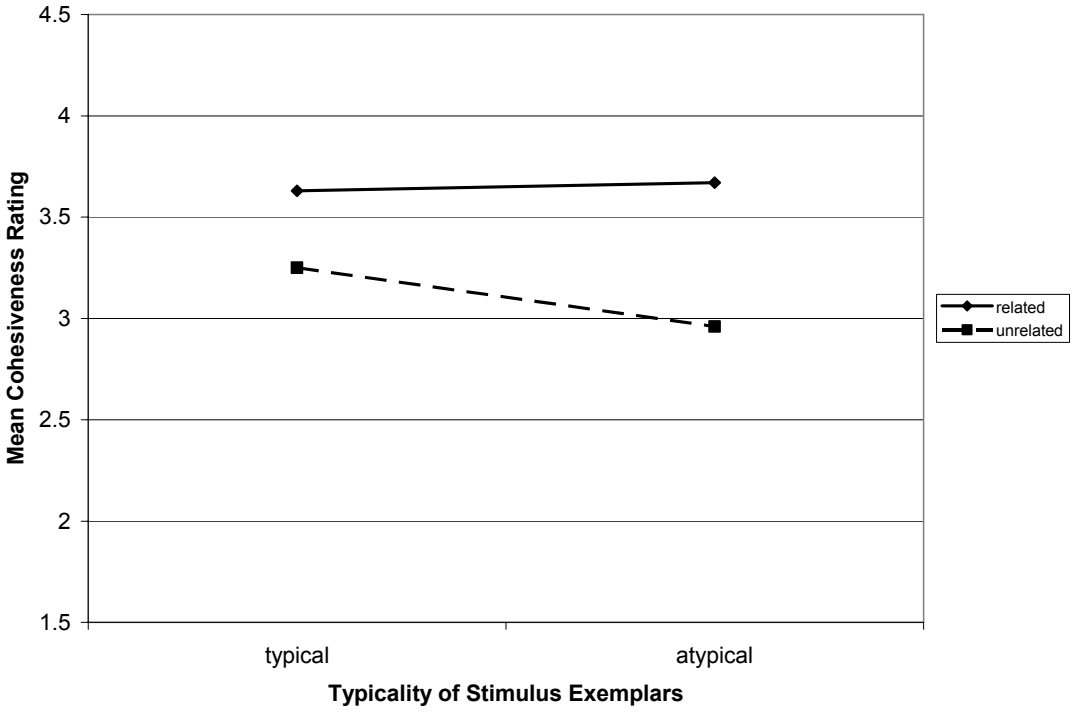

4.5
4
3.5
3
2.5
2
1.5
Mean Originality Ratings
related
unrelated
biological
artifactual

Figure 5.


4.5
4
3.5
3
2.5
2
1.5
Mean Cohesiveness Rating
typical
atypical
Typicality of Stimulus Exemplars
related
unrelated

.

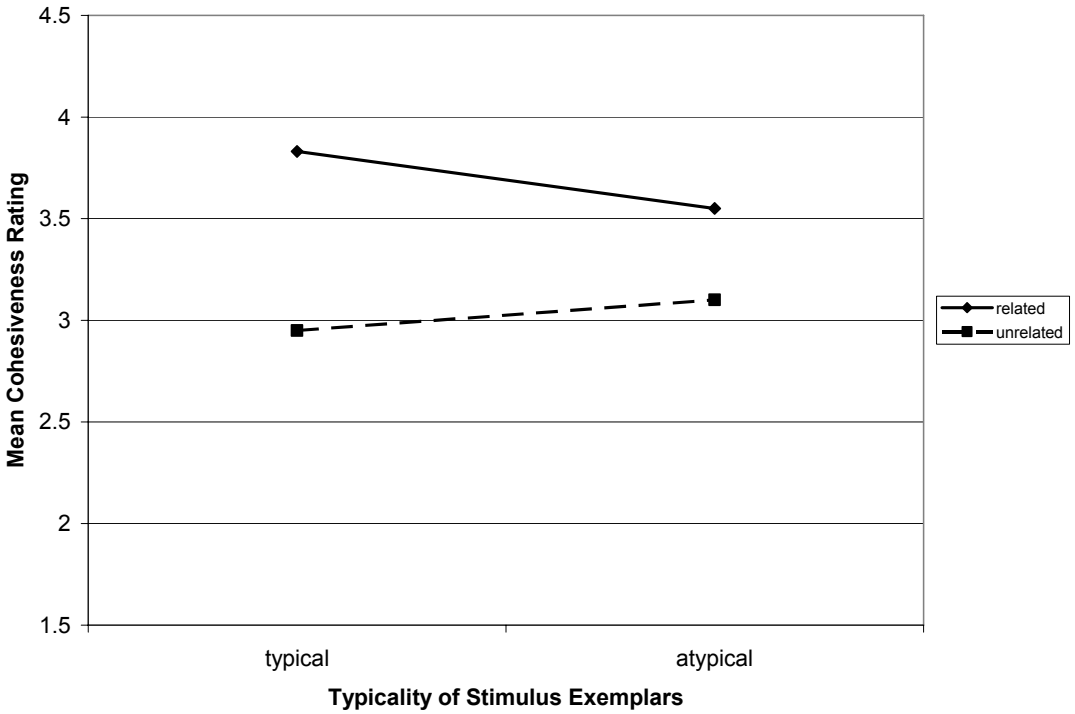

4.5
4
3.5
3
2.5
2
1.5
Mean Cohesiveness Rating
typical
atypical
Typicality of Stimulus Exemplars
related
unrelated

Figure 6.

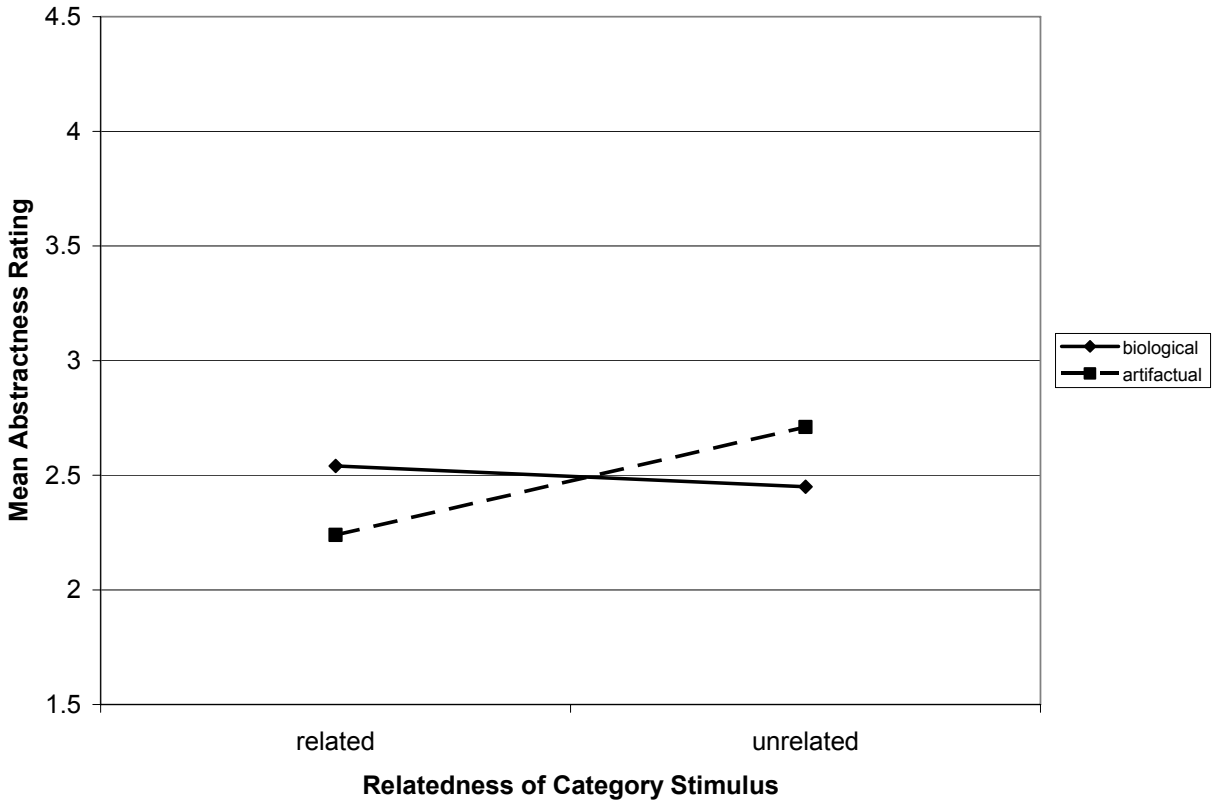
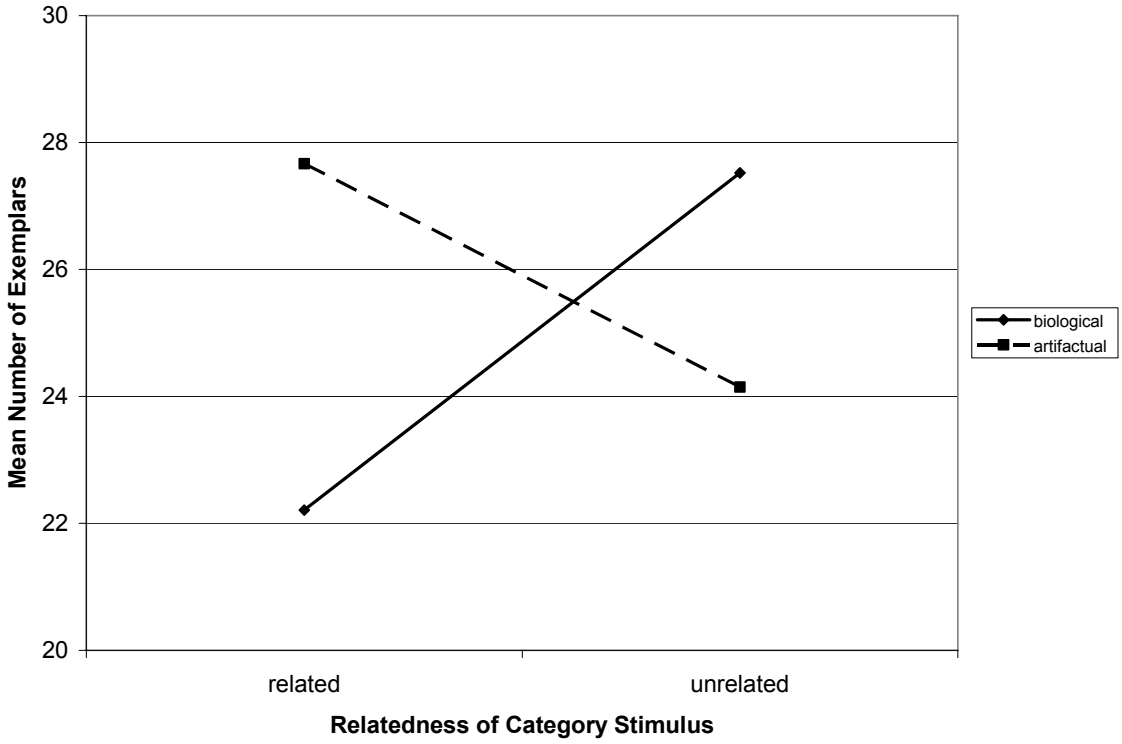

4.5
4
3.5
3
2.5
2
1.5
Mean Abstractness Rating
related
unrelated
Relatedness of Category Stimulus
biological
artifactual

Figure 7.


30
28
26
24
22
20
Mean Number of Exemplars
related
unrelated
Relatedness of Category Stimulus
biological
artifactual