SUBTYPING ENHANCES SUPERORDINATE-LEVEL LEARNING OF DISPERSED CATEGORY STRUCTURES

Alex Gerdom

Submitted to the faculty of the Department of Psychological and Brain Sciences in partial fulfillment of the requirements for the Honors degree in Psychology

INDIANA UNIVERSITY

June, 2016

Accepted by the Psychological and Brain Sciences Faculty, Indiana University, in partial fulfillment of the requirements for the Honors degree in Psychology.

Robert Nosofsky, PhD (Chair)

Robert Goldstone, PhD

John Kruschke, PhD

Date of Oral Examination (June 13, 2016)

Dedicated To Mary Gilbert

Acknowledgements

The preparation of a thesis is not an easy process, but the research and work that have gone into it has taught me more than anything else in my career as an undergraduate. It has not been a solo effort however, and there are many people who helped me in the process. I would like to start by thanking Rob for giving me this opportunity, for all of his guidance, and for entrusting me with the freedom to participate in and learn from each and every stage of the research process. I would like to thank Craig Sanders, Rui Cao, Evelyn Bai, Brian Meagher, Mikaela Akrenius, and everyone else in the lab. I would also like to thank Jordan Delong. I have many fond memories from his independent readings course with Corey Wince and Brenna Williams, and I could not have done any of this without his guidance and support. I would like to thank all my friends, especially Austin Linville, my friend and roommate who was there with me on many of the late nights spent in the process of completing my thesis. Most of all I want to thank my family, for supporting me through it all especially when times were hard. Finally, thanks to everyone who I have no doubt missed here. The start of this document couldn't possibly contain all of you. And though I may have forgotten to mention you here, I will never forget everything you've done to help me in this journey.

Alex Gerdom

Subtyping Enhances Superordinate-Level Learning

of Dispersed Category Structures

Abstract

Geological taxonomies traditionally divide rocks into broad superordinate categories (Igneous, Sedimentary, and Metamorphic) each comprising many subtypes (e.g. Granite, Slate, or Sandstone). We investigated the interaction of category structure and attended category level in learning these categories. In a supervised learning experiment, subjects learned to classify photographs of rocks by either superordinate category or subtype. In a compact condition, similarities between subtypes within each superordinate category were high, whereas in a dispersed condition they were low. Hypothesizing an interaction between the compactness of categories and learned level, an advantage was found for learning of superordinate categories when categories were compact if subjects learned categories at the superordinate level. However when categories were dispersed, learning at the subtype level was found to be advantageous.

Introduction

For a student of the geological sciences, few things can compare to the shift in world-view that the ability to categorize rocks can provide. If one can correctly categorize a rock, then one may make inferences about geological events occurring billions of years in the past. However this classification proves far from trivial. Classification can take place at one or more levels of specificity, with categories nesting and interrelating within a rich hierarchical system. By and large, our understanding of how categories situated in hierarchies such as these are learned is only in its nascent stages, and many questions about the manner in which learning functions in these domains remain either under-explored or unexplored within the category learning literature.

Two features of category hierarchies are critical to the present investigation. First, since such categories can be learned at one or more levels of specificity, attending to learning at one level may potentially influence learning at another. Second, higher level categories can be more or less homogenous depending upon the similarity of their subtypes.

This thesis addresses the intersection of these two factors and will make the somewhat counter-intuitive claim that if one wishes to learn to classify at a superordinate level, it may sometimes be better to learn subtype classifications in addition to the superordinate classifications, rather than the superordinate classifications alone: even if doing so means learning more classifications in the process. Specifically, we hypothesize that the level one should attend to will depend on the structure of the categories one is dealing with. The rest of the Introduction proceeds as follows. First, I will begin by reviewing work on learning of categories at multiple levels of a hierarchy, and recent evidence suggesting that learning at one level may influence learning at others. Then I will introduce the domain of geological categories, which will be used to motivate the problem of coherence of category structures and the role this variable plays in the present work.

Learning at Multiple Levels: Despite great gains in our understanding of the representation and use of categories over the past few decades, relatively little research to date has been done involving category learning including more than one level of specificity. This is not however without exception.

Lassaline, Wisniewski, and Medin (1992) conducted a series of experiments to determine whether level advantages could be obtained in situations where categories lack defining features and to evaluate three categorization models on their ability to predict the pattern of effects. In the study, subjects were shown images of tool-like artificial stimuli that could be classified into two general level categories each of which could be further sub-classified into two specific level categories. Whether general level categories or specific level categories were learned was manipulated between subjects, with subjects learning either at the general level or at the specific level. Subjects were trained on a category verification task in which they were shown a stimulus and a category label, and asked to determine whether the label denoted the correct classification for the object. Subjects were then tested using a speeded version of the category verification task, and models were compared with respect to whether they correctly predicted whether the general level group or the specific level group had a higher probability of correct classification in training. The three models compared were a category utility measure (Gluck & Corter, 1985), the adaptive network model (Gluck & Bower, 1988b), and the exemplar-based context model (Medin & Schaffer, 1978).

In the first of three experiments, stimuli varied along two shape and two texture dimensions. Features varied such that no feature or combination of features were both necessary and sufficient to determine the category of an object. Lassaline et al. found that under these conditions, the general level learning group learned the categories more easily than the specific level learning group: demonstrating level effects can be obtained for fuzzy categories. All three models correctly predicted a general level advantage in this case.

The second of the three experiments was similar to the first in that no feature or combination of features was both necessary and sufficient to predict category membership. The experiment was stricter however, in that no feature was sufficient for determining category membership. Stimuli varied along two shape and one texture dimension. Subjects who learned specific level categories were found to learn categories more easily, and to have faster response times and higher accuracy responses in the speeded verification task used for the test phase. In this experiment, models differed in their predictions: with the category utility measure correctly predicting a specific level advantage, and both the context model and the adaptive network model incorrectly predicting a general level advantage.

In the third and most interesting of the experiments, Lassaline et al. investigated how the distribution of diagnostic features may affect which level is easier to learn. Stimuli varied along two shape dimensions and two texture dimensions. In addition to the between subjects manipulation of learned level (general vs. specific), subjects were additionally assigned to one of two stimulus set conditions (1-Dimensional vs. 4-Dimensional). In the 1-D conditions, diagnostic features fell along a single dimension. In the 4-Dimensional conditions, diagnostic features were spread along four separate dimensions.

In the case where diagnostic features were spread along a single dimension, subjects committed fewer errors on average if they learned categories at the specific level than if they learned categories at the general level. However when features were spread across 4 separate dimensions, the effect was reversed; subjects made fewer errors on average if they had learned general level categories than if they learned at the specific level.

Of the three models surveyed, none of the models managed to correctly predict the interaction between how diagnostic characteristics were distributed and level advantage. The category utility measure correctly predicted a specific level advantage in the 1-D condition, but incorrectly predicted a specific level advantage in the 4-D condition. Both the context model and the adaptive network model predicted general level advantages in both conditions: correctly predicting the 4-D condition, but making incorrect predictions for the 1-D condition.

In summary, the study found that one level or another may be easier to learn even in cases involving fuzzy categories. Additionally, which level is easier to learn may be sensitive to how diagnostic features are distributed across dimensions.

In order to determine if the interaction reported in Experiment 3 of the previous study could be obtained when a more traditional categorization task was used and to see if more recent categorization models could account for this interaction, Palmeri (1999) set out to extend the experiment with three primary differences. First, rather than the tool-like stimuli in the previous study, line drawings of rocket ships varying in wing, nose, tail, and porthole shapes were used. Second, where the preceding study used a category verification task, the extension made use of a forced choice categorization task in which subjects had to correctly choose the category of an item from several possible categories. Third, whereas Lassaline et al. (1992) had compared models on their ability to account for average accuracy in training, this study instead evaluated models on their ability to predict subjects' patterns of performance throughout training. The three models evaluated were the Rational Model (Anderson, 1990), the Configural-Cue Model (Gluck & Bower, 1988a), and ALCOVE (Kruschke, 1992). The latter model is a version of the exemplar-based context model that learns categories on a trial-by-trial basis.

Even though the shift from a category verification task to a forced choice task meant that probability of making a correct categorization by chance was substantially higher in the general level learning conditions than in the specific level learning conditions (50% in the former vs. 25% in the latter), a similar level by distribution effect was observed. In the 1-D condition, subjects learned categorizations more quickly if they learned categories at the specific level than if they learned categories at the general level. However, in the 4-D condition, subjects learned categories more quickly if they learned categories at the general level than if they learned categories at the specific level. Of the three models evaluated, only ALCOVE was able to correctly predict the interaction between learned level and category structure.

Effectively, Palmeri (1999) managed to show that the interaction reported in the preceding study could be replicated under a more traditional category learning paradigm. Additionally, the success of ALCOVE in predicting this interaction demonstrated that models designed to characterize category learning at a single level of abstraction could also manage to account for phenomena where learning functions across multiple levels. We now turn to a recent study, which serves as the primary motivation for the present investigation, the results of which suggest there exist cases in which learning at one level may influence learning at others.

In Noh, Yan, Vendetti, Castel, and Bjork (2014), investigators looked at the interaction between intrinsic value (specifically survival relevancy) and participants' ability to induce categories at two levels of organization. Subjects were shown images of snakes belonging to 6 genera and asked to learn either the genus of the snake (the specific level categorization) or one of two broad level categorizations varying in intrinsic value. In the high-intrinsic value condition the distinction for the broad level categories was between venomous and non-venomous snakes. In the low intrinsic value condition, the labels tropical/non-tropical were substituted for the venomous/non-venomous labels.

Over several training blocks, subjects were shown instances of snakes from each genus and instructed to study either the genus or broad-level classification, with the classification from the level they were instructed to attend to appearing below the image on the left, and classification for the other level appearing to the right in parentheses.

Following training, subjects were then tested on their ability to make both genus and broad level categorizations regardless of the level they were instructed to attend to.

What they found was that: (1) subjects performed better on the level they were instructed to attend to, (2) when tested on genus (specific) level classifications, performance for subjects instructed to attend to the specific level was inhibited for those who saw the high value (venomous/non-venomous) broad level classifications relative to those who saw the low value (tropical/non-tropical) distinctions, and (3) when tested on broad level categories, subjects who were exposed to the high value (venomous/non-venomous) broad level classifications performed better than those who saw low-value labels.

These results are interesting here for two reasons: first they showed that incidental learning can take place at multiple category levels regardless of the level subjects were instructed to attend to. Second, if learning at one level can inhibit intentional learning at another, it raises the question of whether there are cases where learning at one level may enhance learning at another.

As the authors in the preceding study note, several characteristics of snakes could be used to efficiently distinguish between the high level categories in the study. For example: venomous snakes tend to possess arrow-head shaped heads, slit pupils, and have thicker, shorter bodies than non-venomous snakes. (Since in the low-intrinsic value conditions the tropical/non-tropical labels simply replaced the venomous/non-venomous labels, these same characteristics could be used to distinguish between high-level categories in the low intrinsic-value conditions as well.) However there are many situations in which the high-level classifications may be less straightforward. Such scenarios are likely to arise in scientific classifications, where the grounds for a classification may not be based in the visible characteristic features of an object, but rather in knowledge of properties of the object hard-won through scientific investigation. If we now consider a system of categories such as used in geology, we will see a case where the broad level categories have disorganized structures, making it difficult to learn them directly. However, by learning the subtype level categories in addition to the higher-level categories, it may be possible for incidental learning of the high level categories to be superior to direct learning at the superordinate level.

Geological Categories: Though there exist many diverse systems of geological classification, almost invariably, the first system learned by students will be a simplified version of taxonomies used in petrology: a branch of geology concerned with the origin of rocks. At the highest level, these systems will typically classify rocks into three major superordinate categories based on their mode of formation (Marshak, 2012). The first type is Igneous rocks, which are formed from the cooling and solidification of molten rock. The second type is Sedimentary rocks, which are formed through the breakdown of other rocks, via erosion, weathering, and the accumulation of sediment. The final type is Metamorphic rocks, which are formed when events such as heat or pressure cause structural changes to instances of any of the three rock types.

At a lower level, these categories divide into many subtypes. The grounds along which the Igneous, Sedimentary, and Metamorphic types are subdivided is a more complicated affair than that of the initial high-level classification. Though the ways in which these subtypes are classified is complicated and beyond the scope of the present thesis, the names of some of these subtypes may be familiar to the reader. For instance, Granite is one subtype of Igneous rock, Sandstone is a subtype of Sedimentary rock, and Slate is a subtype of Sedimentary rock. A fuller set of examples is provided in Figure 1.

Compact and Dispersed Category Structures: With geological categories in tow, we now have a concrete example with which we can motivate the issue of the coherence of category structures (e.g., Richler & Palmeri, 2014). Category structures may vary in their level of dispersion, which for the purpose of this study concerns to what degree the category subtypes cohere together to form discrete regions in a feature space.

In order to illustrate this principle, let us assume that we have 3 Igneous, 3 Sedimentary, and 3 Metamorphic subtypes. Depending upon how similar subtypes from each category are to one another, the category structure we encounter will fall somewhere along a spectrum ranging from highly compact to highly dispersed. These two extremes are illustrated in Figure 2. At the compact end of this spectrum, we have a case where the subtypes that make up a superordinate category are all more similar to subtypes of the same superordinate category than they are to subtypes of contrasting superordinate categories. At the dispersed end of this spectrum, we have a case where the subtypes that make up a superordinate category are all more similar to subtypes of other superordinate categories than they are to subtypes of their own superordinate category.

That dispersed structures of this type exist and may be common should not surprise us. For example in the case of geology, since the categories of Igneous, Metamorphic, and Sedimentary rocks are defined by how they formed rather than any particular features they possess, we should have no particular reason to expect various instantiations of them to resemble one another.

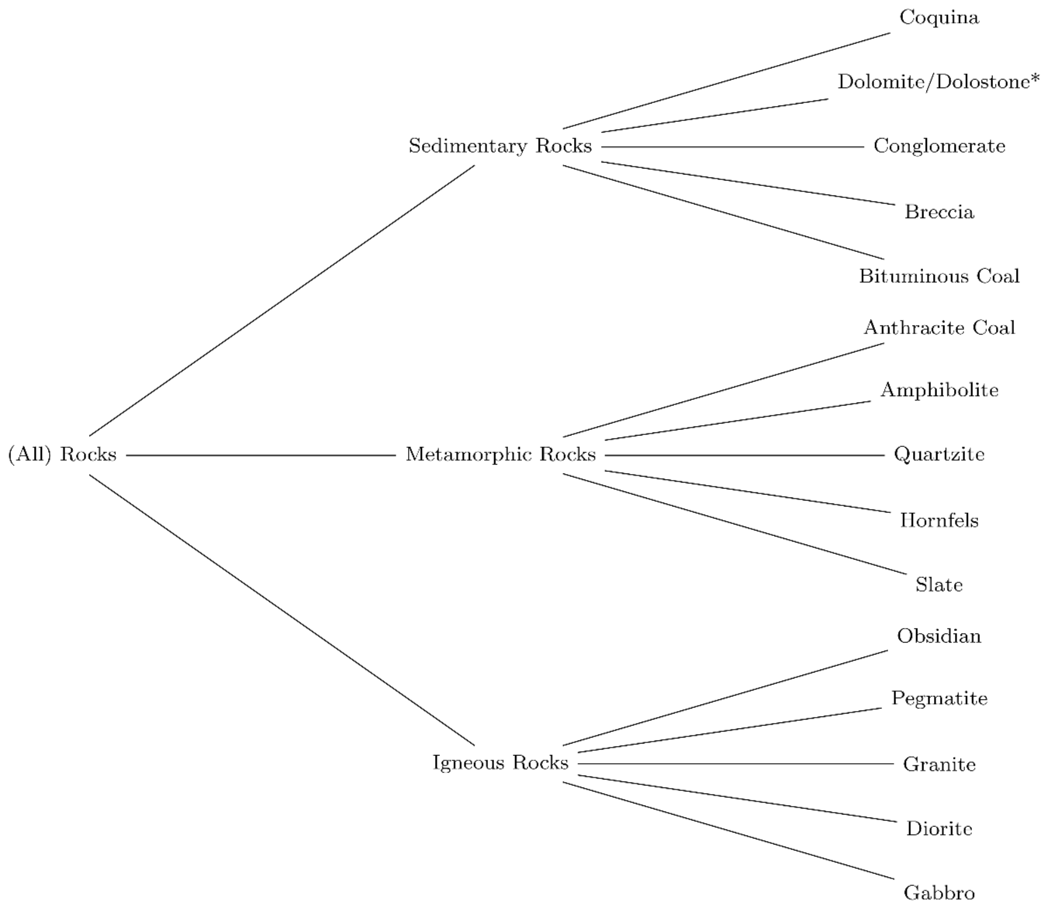

Figure 1: Schematic illustration of a hierarchy of rock types. The hierarchy is drawn with generality decreasing from left to right. From left to right it is ordered under class inclusion. From right to left, the taxonomy is characterized by “is a type of” relations. *Not to be confused with the mineral also known as dolomite.

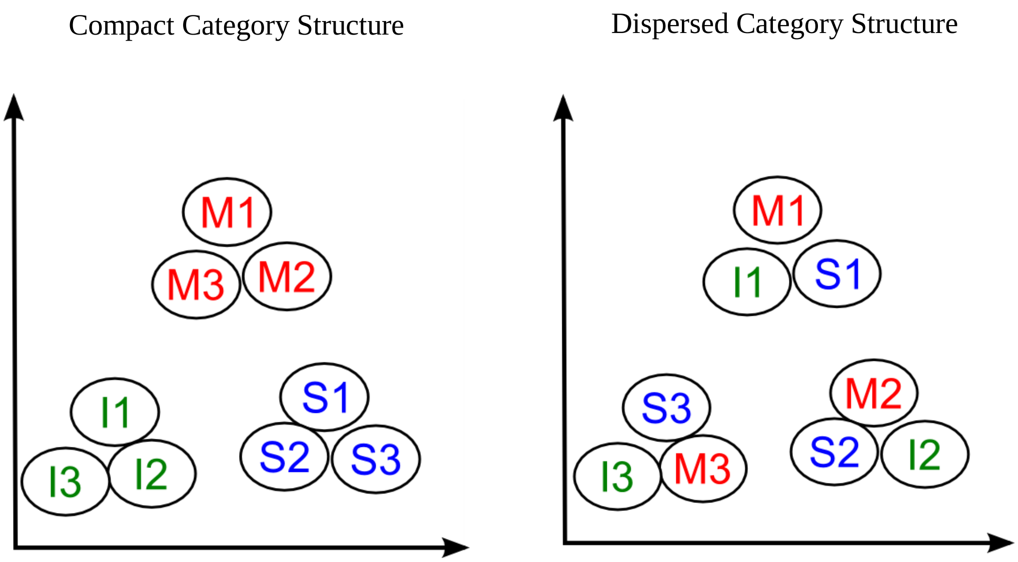

Figure 2: Schematic illustration of two types of category structures (where I = Igneous, M = Metamorphic, S = Sedimentary). "M1" stands for Metamorphic Subtype 1, and so forth. Categories grouped closer together are perceived as more similar to one another than categories farther away. In the compact category structure, all subtypes of each category are more similar to one another than to subtypes of other superordinate categories. In a dispersed category structure, subtypes of each category may be more similar to subtypes of other superordinate categories than to other subtypes of the same superordinate category.

We hypothesized that, for compact structures, it will be advantageous to learn the high-level categories directly; whereas for dispersed structures, learning the subtypes may be advantageous for high-level learning. Our reasoning behind this hypothesis is that compact category structures have good “signal-to-noise” ratio (high within-category similarity and low between-category similarity), so it is optimal to focus on this level. However as categories grow more dispersed, the signal-to-noise ratio is greatly reduced. The suggested remedy for this problem is to subtype the superordinate categories. By subtyping and learning the superordinate categories indirectly, subjects may be able to avoid issues that arise with disorganized high-level category structures, because the signal-to-noise ratio for any given category is higher at the subtype level.

In order to evaluate this hypothesis, using real world materials drawn from geological classification, we designed a supervised category learning experiment to investigate the potential interaction between the compactness of category structures and learned level. Due to lack of basic research in the area, any result will help to greatly increase our understanding of category learning involving multiple levels of abstraction. Additionally, if learning of difficult classifications such as superordinate rock types can be enhanced by learning of the category subtypes, then such a finding may help in the development of new educational techniques for teaching those categories.

Methods

Subjects: 132 undergraduate students took part in a supervised category learning experiment. Participants were enrolled in introductory psychology courses at Indiana University and agreed to participate as part of fulfillment of an experiment participation requirement. Each participant was assigned to one of four conditions using a systematic random sampling technique. To simplify issues of interpretation that arise in mixed-factors designs with unbalanced cells, the analysis discussed in this section was restricted to the first 30 subjects from each condition.

Materials: Materials consisted of images of rocks gathered from online repositories (Geology.com, SandAtlas.org, Geoscience Digital Image Library (GeoDIL), GeoScenic Portal), with minor editing performed using PAINT.net photo-editing software to remove potentially distracting features such as text, keystoning, and physical labels used for archival purposes. The images were used to construct two sets of stimuli, each with 9 subtypes represented (3 igneous, 3 sedimentary, 3 metamorphic) with 6 images per subtype. Stimuli where shown on 3 PCs in the lab using the Psychophysics Toolbox extensions for MATLAB (Brainard, 1997; Kleiner et al., 2007). The subtypes in each set of stimuli are shown in Tables 1 and 2. Subtypes were chosen to produce sets with compact or dispersed structure as discussed in the Introduction. Category structures were verified in a subsequent similarity scaling experiment which used pairwise similarity ratings for the full combined set of images to compute a multidimensional scaling solution (Shepard, 1980). A 3-dimensional solution provided a good first-order account of the data, with the 3 dimensions characterized by lightness (how dark or light in color a rock was), average grain size, and “sorting” (how much there existed a mix of small and large grain sizes for rocks that had a visible grain). Illustrations of the scaling solution are provided in the set of figures in the appendix.

Dispersion statistics for each stimulus set are summarized in Table 3. The average within-category similarity and the average minimum between-category similarity were calculated for each superordinate category. The average within-category similarity was

Table 1: Stimuli for Compact Condition

Labels		Distances		Info	
Category	Subtype	$\bar{d}_{within}$	$d_{minbtwn}$	Name	Example
"Igneous"	"Igneous 1"	0.522	0.465	Gabbro	
"Igneous"	"Igneous 2"	0.485	0.626	Granite	
"Igneous"	"Igneous 3"	0.380	0.657	Diorite	
"Metamorphic"	"Metamorphic 4"	0.452	0.798	Slate	
"Metamorphic"	"Metamorphic 5"	0.478	0.657	Quartzite	
"Metamorphic"	"Metamorphic 6"	0.468	0.465	Hornfels	
"Sedimentary"	"Sedimentary 7"	0.545	0.928	Conglomerate	
"Sedimentary"	"Sedimentary 8"	0.682	0.758	Coquina Limestone	
"Sedimentary"	"Sedimentary 9"	0.621	0.626	Breccia	

Table 2: Stimuli for Dispersed Condition

Labels		Distances		Info	
Category	Subtype	$\bar{d}_{within}$	d_{minbtwn}	Name	Example
"Igneous"	"Igneous 1"	0.920	0.390	Pegmatite	
"Igneous"	"Igneous 2"	1.086	0.626	Granite	
"Igneous"	"Igneous 3"	1.114	0.068	Obsidian	
"Metamorphic"	"Metamorphic 4"	0.781	0.587	Amphibolite	
"Metamorphic"	"Metamorphic 5"	0.957	0.462	Quartzite	
"Metamorphic"	"Metamorphic 6"	0.858	0.068	Anthracite Coal	
"Sedimentary"	"Sedimentary 7"	1.048	0.462	Dolomite (Dolostone)	
"Sedimentary"	"Sedimentary 8"	1.104	0.168	Bituminous Coal	
"Sedimentary"	"Sedimentary 9"	1.056	0.390	Breccia	

Table 3: Dispersion Statistics

	Compact Set		Dispersed Set	
Category	$Mean(\bar{d}_{within})$	$Mean(d_{minbtwn})$	$Mean(\bar{d}_{within})$	$Mean(d_{minbtwn})$
Igneous	0.462064	0.582931	1.040096	0.361470
Metamorphic	0.466103	0.640262	0.865311	0.372043
Sedimentary	0.615845	0.771006	1.069693	0.340187

calculated by first calculating the mean distance for each subtype to the other two subtypes of the same superordinate category ($\bar{d}_{within}$), and then averaging these distances across the three subtypes of each superordinate category. The average minimum between-category similarity was calculated by first, for each subtype, finding the distance to most similar subtype belonging to a contrasting superordinate category ($d_{minbtwn}$), and then averaging these distances across the three subtypes of each superordinate category. (The complete distance matrices are available in the appendix.) The statistics for both the compact and dispersed sets were found to accord with our subjective impressions regarding the category structures: In the compact set, average distances to other subtypes within a superordinate category were lower for each subtype than were minimum distances to subtypes of other superordinate categories. In the dispersed set, the minimum distance to a subtype of another superordinate category was lower for any given subtype than was average distance to other subtypes of the same superordinate category in all but one case.

Procedure: Subjects were assigned to one of four conditions, corresponding to two levels of learning level (direct, high-level learning or indirect, subtype-level learning), and two levels of stimulus set dispersion (compact or dispersed).

At the onset of the experiment, a screen was presented instructing participants how to input their answers. In the superordinate learning condition, subjects used the "I", "S", and "M" keys to categorize rocks as Igneous Sedimentary, or Metamorphic, respectively. In the sub-type learning condition, subjects used keys 1-9 at the top of the keyboard, which were labeled with the high-level category and subtype number (e.g., “Sedimentary 7").

For each subject, 3 images from each subtype were randomly selected to serve as training examples, and the remaining 3 images from each subtype were retained for use in a transfer phase in which all images were shown.

The training phase of the experiment consisted of 3 blocks of 54 trials where feedback was provided. During these blocks, each of the three training tokens from each subtype was presented twice. The order of presentation of the 54 stimuli was randomized for each subject. On each trial, subjects were supplied with a stimulus screen, such as that in Figure 3, asking them to classify the rock into a category. Upon supplying a response, the text under the image changed to say "Correct!" if the response was correct, otherwise it would change to say "Incorrect!" followed by the correct category. In the high-level learning condition, the feedback was with respect to only the high level of categorization (e.g., “Incorrect! Igneous”), whereas in the subtype-learning condition, the feedback was with respect to the subtype level (e.g., "Incorrect! Igneous 2"). At the end of each block subjects were informed by the computer of their overall percentage correct for that block. In the direct high-level learning condition, this meant subjects were shown the percent of trials on which they provided a correct superordinate category response. In the subtype-learning conditions, this meant that subjects were shown the percent of trials for which they had provided the correct subtype level response.

The fourth and final block was a transfer phase consisting of 108 trials without feedback. During this block, both the 27 old training tokens as well as 27 new transfer items were presented. Each token was presented twice, again in a random order for each subject. The same prompt was used as in the training blocks, however when a response was supplied the text changed to say "Okay!"

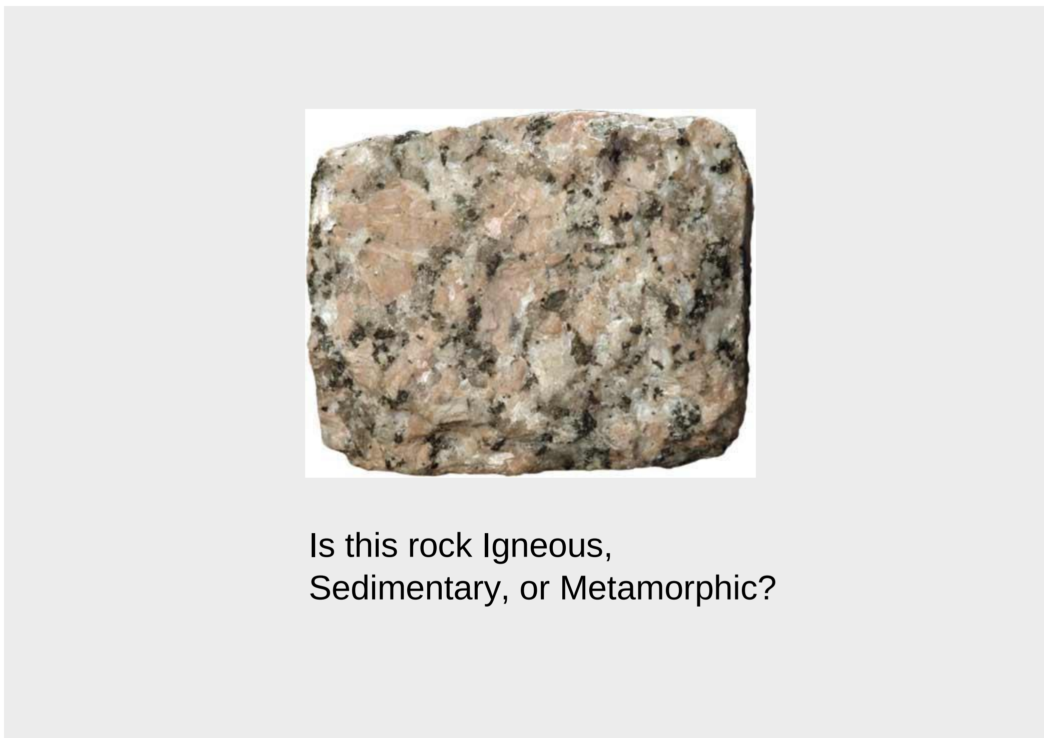

Figure 3: Example of the stimuli and response prompts used in the experiment. Image dimensions: 300x300px. Background color: (R:100, G:100, B:175)

Analysis: Two measures were taken on the transfer phase of the experiment. The first measure was the percent correct for stimuli that had appeared during training with respect to the superordinate category. The second measure was percent correct for novel stimuli which appeared only in the transfer phase, also with respect to the superordinate category. Note that while subjects in subtype learning conditions were shown their percent correct in terms of correct subtype responses, we actually analyzed the percentage of trials with responses belonging to the correct superordinate category. This method of scoring is sensible because our hypotheses pertain to how well subjects will classify at the high (superordinate) level of classification. Using these measures a 2x2x2 mixed factors analysis of variance was performed using the ezANOVA package for the R programming language. Between-subjects factors were learning level (Direct-superordinate learning, Subtype learning) and category structure (Compact, Dispersed). The within subjects factor was whether the stimuli were seen previously during the training phase or were novel and shown only in the transfer phase.

Results

Performance for each group during the transfer phase is presented in Figure 4. Main effects were observed for Stimulus Novelty [F (1,120) = 384.0, p < .001, η_G^2=0.393], with subjects performing better on stimuli seen during training than on novel stimuli; as well as for Category Structure [F (1,120) = 182.0, p < .001, $\eta_G^2 = 0.547$] with participants who learned the compact set performing better than those who received the dispersed set.

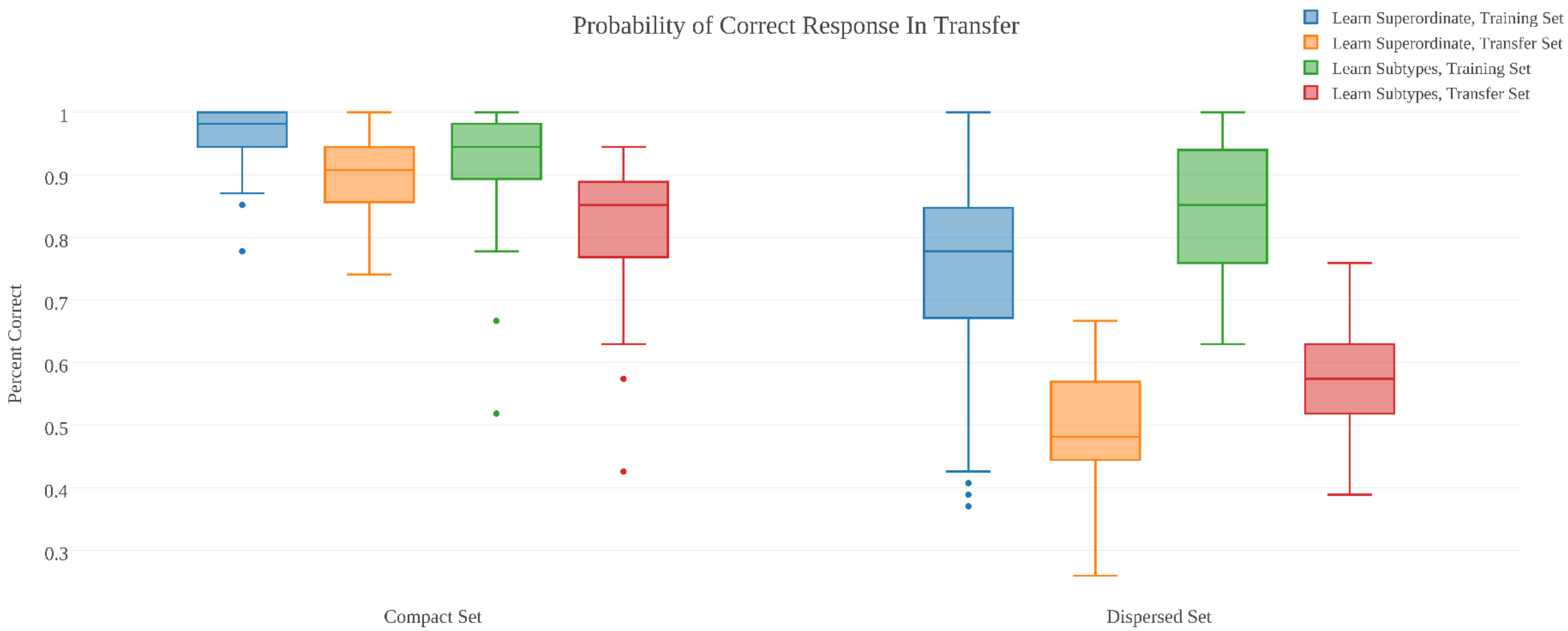

Figure 4: Probability of correct response with regards to superordinate classification.

The striking result however, was a strong interaction between Category Structure and Learned Level [F (1,120) = 18.6, p < .001, $\eta_G^2 = 0.11$]. Subjects learning the compact set performed better if they learned the superordinate categories directly. However if they learned the dispersed set the direction of the effect reversed: subjects learning the dispersed set performed better if they learned the superordinate categories indirectly by also learning the subtypes than if they were only focused on learning the superordinate categories. Finally there occurred an interaction between Category Structure and Stimulus Novelty [F (1,120) = 98.7, p < .001, $\eta_G^2 = 0.143$], with greater differences between performance on training stimuli and performance on transfer stimuli in the dispersed condition than in the compact condition.

Discussion

We began with the question: if one's goal is to learn categories at the superordinate level, is the better strategy to learn the superordinate categories alone or should one attempt to simultaneously learn at the subtype level as well? We had hypothesized that the answer would depend upon the structure of the superordinate categories being learned. Specifically, when compact, then direct high level learning of the superordinate categories would lead to better performance than simultaneous subtype learning; but when dispersed, then simultaneous subtype learning would allow for better classification than learning the superordinate categories alone. Our results provided strong support for our hypotheses. Indeed, the predicted interaction was observed both for images shown during training, and for novel generalization images for which feedback was never provided.

In the category-learning literature it is generally assumed that learning fine-grained distinctions that are not relevant will subtract from one's ability to learn the relevant high-level classifications. Our results show that this assumption may not always be true. Breaking up a category into a number of subtypes and learning those subtypes in addition to the superordinate category may in fact improve learning outcomes if the set of categories one is dealing with has a sufficiently dispersed category structure.

These findings are important for several reasons: First, to the best of my knowledge, this may be one of the first documented cases of indirect learning of higher level categories managing to surpass direct high level learning of those same categories. Second, this result contributes to a small list of studied phenomena involving learning of categories from multiple levels in a category hierarchy. It is my hope that results such as this will help to spur increased research interest in this long neglected area. Third, that this result was obtained using complex real world categories drawn from the natural sciences, gives us reason to hope that this result may go on to help serve as part of a basis for future research aimed at developing new educational strategies for teaching such classifications. We turn now to factors limiting our understanding of how these results may generalize and provide suggestions for future work.

Perhaps the most pressing limitation to our understanding of how these results generalize is that it is not immediately obvious what role was played by the nomenclature used for the labels in our study. In the present study, subtypes were labeled with the superordinate category along with an identifying subtype number (e.g. "Igneous 1"). However there are many contexts in which it would be preferable to actually use the name of the subtype. Replication of the present study using the actual names of the subtypes would help to increase generalizability for use in the design of classroom materials, where subtypes may be interesting both in their own regard and in how they inform students' understandings of their superordinate categories. More generally, development of a better understanding of how superordinate and subtype labels come to be associated in various learning contexts could be highly productive theoretically and practically.

Considerably more work needs to be done to determine to what extent categories in the natural world tend to display compact or dispersed structure. Ultimately this is a question that will have to be answered in the course of research over the long term. Perhaps the best source of evidence will have to come through similarity scaling studies of categories drawn from fields such as geology, biology, and other natural sciences to understand what kind of structures arise in these fields.

Finally, while we have shown that learning subtypes in addition to superordinate categories can enhance superordinate level classification in dispersed category domains, we still do not have a deep theoretical explanation for why this result arises. Future research is needed to answer this question.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Psychology Press.

David H, B. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Gluck, M. A., & Bower, G. H. (1988). Evaluating an adaptive network model of human learning. *Journal of Memory and Language*, *27*(2), 166–195. http://doi.org/10.1016/0749-596X(88)90072-1

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, *117*(3), 227–247. http://doi.org/10.1037/0096-3445.117.3.227

Gluck, M. A., & Corter, J. (1985). Information, uncertainty, and the utility of categories. In Proceedings of the Seventh Annual Conference of the Cognitive Science Society (pp. 283--288).

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.

Kruschke, J. K. (1992). ALCOVE: an exemplar-based connectionist model of category learning. *Psychological Review*, *99*(1), 22–44. http://doi.org/10.1037/0033-295X.99.1.22

Lassaline, M. E., Wisniewski, E. J., & Medin, D. L. (1992). Basic levels in artificial and natural cateogries: Are all basic levels created equal? In B. Burns (Ed.), *Percepts, concepts and categories* (Vol. 93, pp. 327–378). North-Holland. Retrieved from http://www.sciencedirect.com/science/article/pii/S0166411508610134

Marshak, S. (2012). *Essentials of Geology (Fourth Edition)* (4th ed.). W. W. Norton & Company. Retrieved from http://amazon.com/o/ASIN/0393919390/

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*(3), 207.

Noh, S. M., Yan, V. X., Vendetti, M. S., Castel, A. D., & Bjork, R. a. (2014). Multilevel Induction of Categories: Venomous Snakes Hijack the Learning of Lower Category Levels. *Psychological Science*, *25*(8), 1592–1599. http://doi.org/10.1177/0956797614535938

Palmeri, T. J. (1999). Learning categories at different hierarchical levels: a comparison of category learning models. *Psychonomic Bulletin & Review*, *6*(3), 495–503. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12198789

Richler, J. J., & Palmeri, T. J. (2014). Visual category learning. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(1), 75–94.

Shepard, R. N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science*. http://doi.org/10.1126/science.210.4468.390

Table 4: Compact Condition

Category	$d_{minbtwn}$	$\bar{\bar{d}}_{within}$	Name	Subtype	I1	I2	I3	M4	M5	M6	S7	S8	S9
Igneous	0.465	0.522	Gabbro	Ign.1	0	0.627	0.417	0.798	0.696	0.465	1.045	1.163	0.760
Igneous	0.626	0.485	Granite	Ign.2	0.627	0	0.343	1.074	0.717	0.896	0.967	0.821	0.626
Igneous	0.657	0.380	Diorite	Ign.3	0.417	0.343	0	0.876	0.657	0.721	1.146	1.054	0.824
Metamorphic	0.798	0.452	Slate	Met.4	0.798	1.074	0.876	0	0.462	0.442	1.28	1.205	1.272
Metamorphic	0.657	0.478	Quartzite	Met.5	0.696	0.717	0.657	0.462	0	0.494	0.928	0.758	0.916
Metamorphic	0.465	0.468	Hornfels	Met.6	0.465	0.896	0.721	0.442	0.494	0	1.001	1.112	0.937
Sedimentary	0.928	0.545	Conglomerate	Sed.7	1.045	0.967	1.146	1.280	0.928	1.001	0	0.605	0.484
Sedimentary	0.758	0.682	Coquina	Sed.8	1.163	0.821	1.054	1.205	0.758	1.112	0.605	0	0.758
Sedimentary	0.626	0.621	Breccia	Sed.9	0.760	0.626	0.824	1.272	0.916	0.937	0.484	0.758	0

Table 5: Dispersed Condition

Category	$d_{minbtwn}$	$\bar{d}_{within}$	Name	Subtype	I1	I2	I3	M4	M5	M6	S7	S8	S9
Igneous	0.390	0.920	Pegmatite	Ign.1	0.000	0.892	0.948	0.587	1.023	0.920	1.074	0.911	0.390
Igneous	0.626	1.086	Granite	Ign.2	0.892	0.000	1.280	0.644	0.717	1.223	1.046	1.112	0.626
Igneous	0.068	1.114	Obsidian	Ign.3	0.948	1.280	0.000	0.723	1.102	0.068	1.285	0.234	1.212
Metamorphic	0.587	0.781	Amphibolite	Met.4	0.587	0.644	0.723	0.000	0.880	0.682	1.167	0.637	0.659
Metamorphic	0.462	0.957	Quartzite	Met.5	1.023	0.717	1.102	0.880	0.000	1.035	0.462	0.878	0.916
Metamorphic	0.068	0.858	Anthracite Coal	Met.6	0.920	1.223	0.068	0.682	1.035	0.000	1.227	0.168	1.168
Sedimentary	0.462	1.048	Dolomite	Sed.7	1.074	1.046	1.285	1.167	0.462	1.227	0.000	1.097	1.000
Sedimentary	0.168	1.104	Bituminous Coal	Sed.8	0.911	1.112	0.234	0.637	0.878	0.168	1.097	0.000	1.112
Sedimentary	0.390	1.056	Breccia	Sed.9	0.390	0.626	1.212	0.659	0.916	1.168	1.000	1.112	0.000

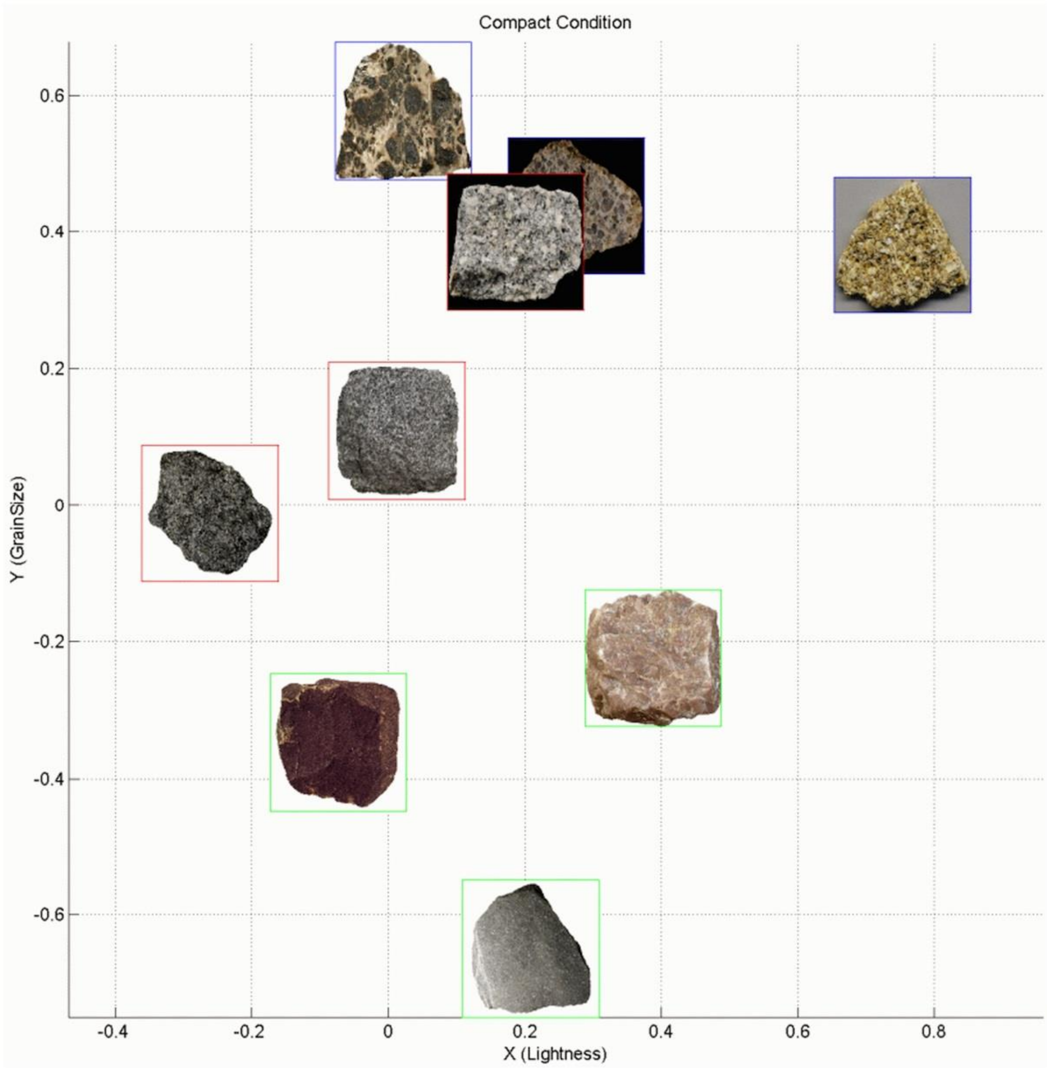

Figure 5: Similarity scaling results for subtypes used in the compact condition of this study (1 of 3)

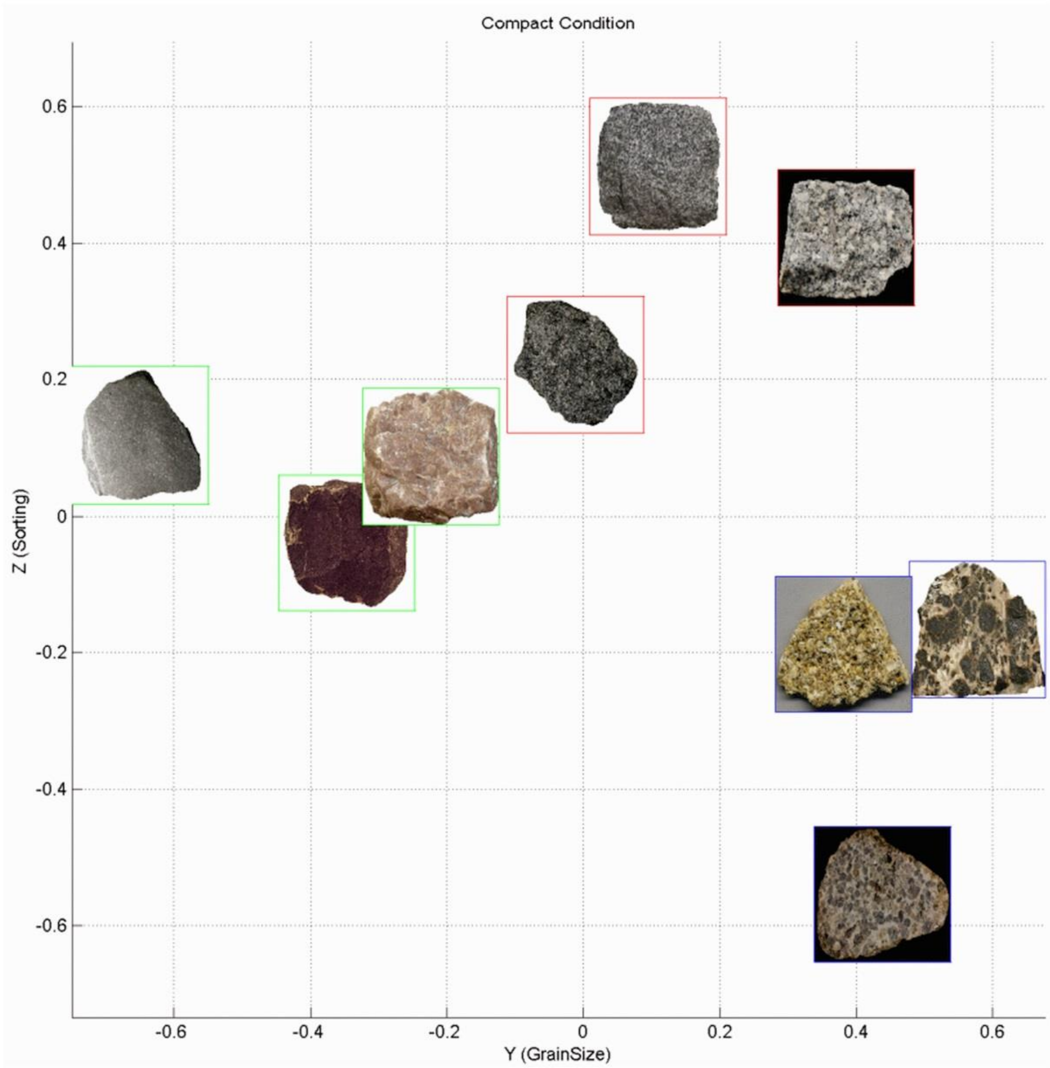

Figure 6: Similarity scaling results for subtypes used in the compact condition of this study (2 of 3)

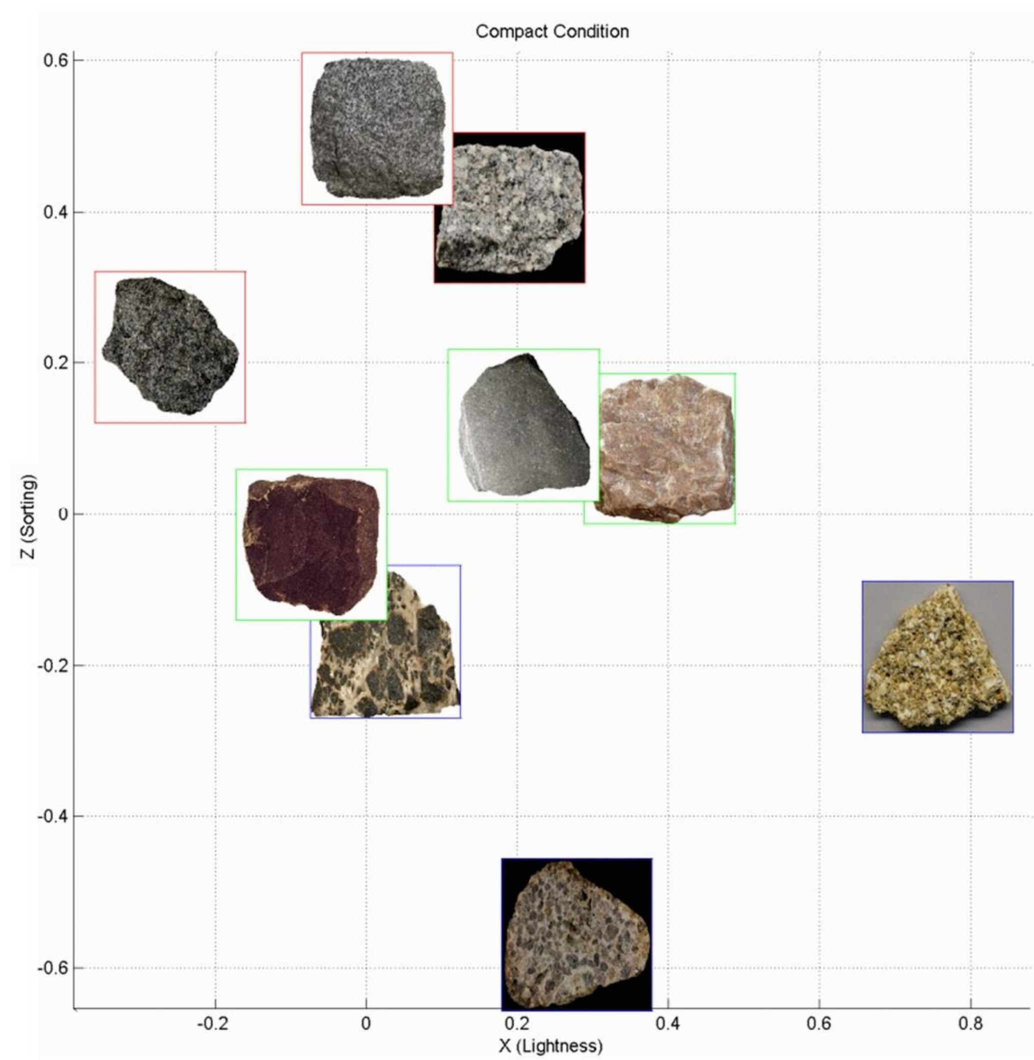

Figure 7: Similarity scaling results for subtypes used in the compact condition of this study (3 of 3)

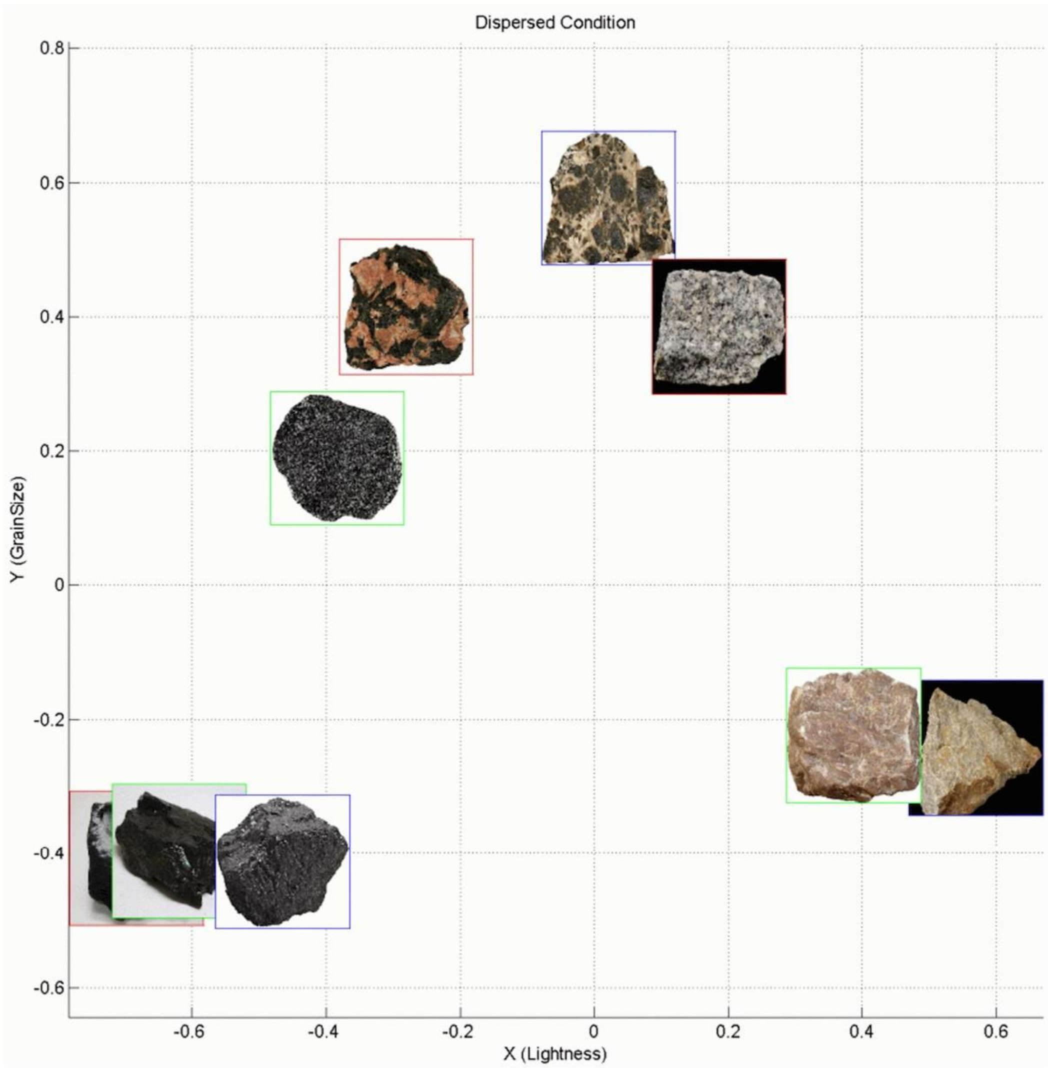

Figure 8: Similarity scaling results for subtypes used in the dispersed condition of this study (1 of 3)

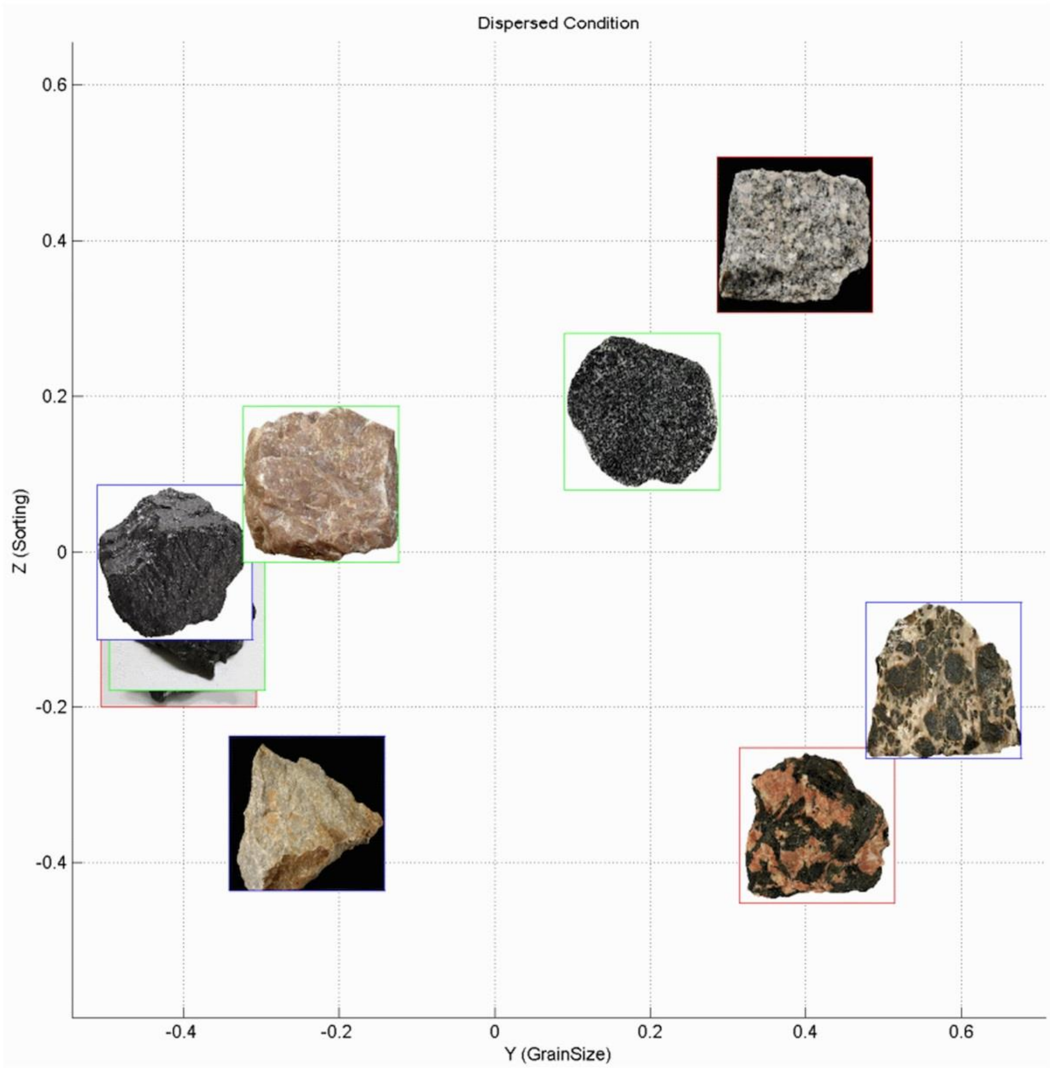

Figure 9: Similarity scaling results for subtypes used in the dispersed condition of this study (2 of 3)

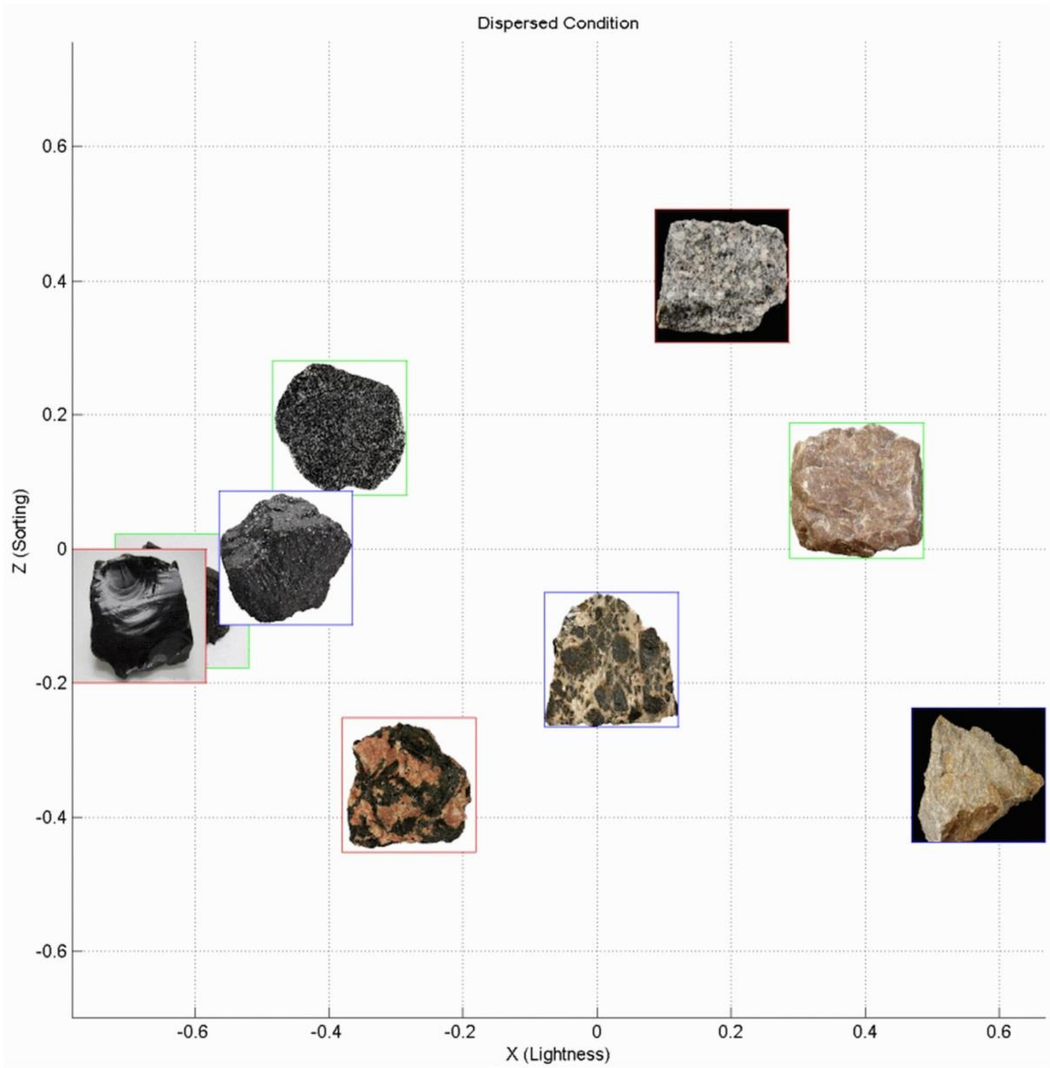

Figure 10: Similarity scaling results for subtypes used in the dispersed condition of this study (3 of 3)