Cogn Process (2018) 19:1–16
https://doi.org/10.1007/s10339-017-0840-7

RESEARCH REPORT

On the learning difficulty of visual and auditory modal concepts: Evidence for a single processing system

Ronaldo Vigo[1] · Karina-Mikayla C. Doan[1] · Charles A. Doan[1] · Shannon Pinegar[2]

Received: 30 September 2016 / Accepted: 11 October 2017 / Published online: 26 October 2017

Abstract The logic operators (e.g., "and," "or," "if, then") play a fundamental role in concept formation, syntactic construction, semantic expression, and deductive reasoning. In spite of this very general and basic role, there are relatively few studies in the literature that focus on their conceptual nature. In the current investigation, we examine, for the first time, the learning difficulty experienced by observers in classifying members belonging to these primitive "modal concepts" instantiated with sets of acoustic and visual stimuli. We report results from two categorization experiments that suggest the acquisition of acoustic and visual modal concepts is achieved by the same general cognitive mechanism. Additionally, we attempt to account for these results with two models of concept learning difficulty: the generalized invariance structure theory model (Vigo in Cognition 129(1):138–162, 2013, Mathematical principles of human conceptual behavior, Routledge, New York, 2014) and the generalized context model (Nosofsky in J Exp Psychol Learn Mem Cogn 10(1):104–114, 1984, J Exp Psychol 115(1):39–57, 1986).

Keywords Concept learning · Auditory concepts · Logic operators · Categorization behavior

Handling editor: Kenneth Kurtz (Binghamton University); Reviewers: John Paul Minda (Western University, Canada), Elliott Moreton (University of North Carolina Chapel Hill), Daniel Little (University of Melbourne).

Electronic supplementary material The online version of this article (doi:10.1007/s10339-017-0840-7) contains supplementary material, which is available to authorized users.

✉ Ronaldo Vigo
vigo@ohio.edu

Karina-Mikayla C. Doan
kb016011@ohio.edu

Charles A. Doan
cd634011@ohio.edu

Shannon Pinegar
spinega@gmail.com

1 Consortium for the Advancement of Cognitive Science, Department of Psychology, Ohio University, 200 Porter Hall, Athens, OH 45701, USA

2 Department of Psychology, Baruch College, One Bernard Baruch Way, New York, NY 10010, USA

Introduction

The ability to learn concepts (i.e., to generalize) plays a fundamental role in making sense of the world around us. A small number of concepts, such as "and," "or," and "not," are fundamental to cognition and are referred to as logical connectives or logic operators. These concepts are at the heart of deductive reasoning and give humans the ability to construct complex compound statements that specify possible propositional alternatives (Byrne and Johnson-Laird 2009; Goodwin and Johnson-Laird 2011; Khemlani et al. 2014; Newstead et al. 1984). For example, a parent may say to a child, "Please pick up your toys *and* make your bed *or* you are *not* playing outside today." Although relatively simple, children learn that in order to play outside they must perform both of the actions in the first clause of the sentence (an instance of the operator *conjunction*). Conversely, if the child has learned the meaning of the *disjunction* operator ("or") and does not want to play outside, then the child may choose to ignore one or both of the parent's requests (much to the parent's chagrin).

Much of our daily lives involve parsing similarly constructed statements, with each statement often instantiating

Table 1 Sixteen Boolean logic operators, truth-table values, and real-world examples

Family type	Logical operator	TT	TF	FT	FF	Real-world examples
$2_2[0]$	Empty set	−	−	−	−	No qualifying members based on given two-dimensional criteria
$2_2[1]$	Conjunction	+	−	−	−	Airport security: must have both a valid I.D. and a boarding pass
	Exclusion (P1, ~ P2)	−	+	−	−	After surgery: do get plenty of rest; do not engage in strenuous exercise
	Exclusion (~ P1, P2)	−	−	+	−	Nutrition advice: do not consume refined sugar; do consume fruits
	Joint denial	−	−	−	+	Safe driving: neither exceed the speed limit nor use a cell phone
$2_2[2]$-I	Affirmation (P1)	+	+	−	−	Demographics: all employed, regardless of whether a degree is held
	Affirmation (P2)	+	−	+	−	Demographics: all degree holders, regardless of employment status
	Absence (P2)	−	+	−	+	Demographics: all non-degree holders, regardless of employment status
	Absence (P1)	−	−	+	+	Demographics: all unemployed, regardless of whether a degree is held
$2_2[2]$-II	Biconditional	+	−	−	+	Voting: you are eligible to vote if and only if you are registered
	Exclusive disjunction	−	+	+	−	Directions: one must turn either left or right when a one-way road ends
$2_2[3]$	Inclusive disjunction	+	+	+	−	Scheduling: my schedule requires we meet on the 3rd and/or 6th of the month
	Conditional (P2 → P1)	+	+	−	+	Groceries: if you buy cereal, then you must also buy milk
	Conditional (P1 → P2)	+	−	+	+	Workplace: you must ask for a raise only if you perform well
	Alternative denial	−	+	+	+	Cinema: two films are showing simultaneously; choose one or neither
$2_2[4]$	Full set	+	+	+	+	All qualifying members based on given two-dimensional criteria

multiple types of logic operators. But, how do observers acquire and learn to use logic operators? And, are some logic operators more primitive to conceptual processing than others? Many researchers interested in categorization and concept learning behavior pursued answers to these questions during the middle of the twentieth century. Following in their tradition, we report findings from two categorization experiments conducted to test fundamental hypotheses regarding the structure of the concepts corresponding to the logic operators or, equivalently, the "modal concepts" (Vigo 2009, 2014; Vigo and Allen 2009). Additionally, we empirically examine whether acquiring and learning to use modal concepts depends on the nature of the incoming sensory information (e.g., visual or acoustic) and we conclude with resulting implications and limitations.

Prior research on the Boolean logic operators

Early research into human categorization and concept learning investigated how observers acquire subsets of the 16 fundamental logic operators belonging to the propositional calculus consisting of two binary-valued stimulus dimensions, also referred to as Boolean operators (Bruner et al. 1956; Bourne 1970; Bourne and Guy 1968; Conant and Trabasso 1964; Haygood and Bourne 1965; Neisser and Weene 1962; Vigo 2009, 2014; Vigo and Allen 2009). This is hardly surprising given that we use Boolean operators everyday in normal discourse. Take, for example, the phrase, "Must have both a valid I.D. and a boarding pass" (TT; see Table 1). This is a common instance of the operator *conjunction* ("and") where two properties have attributes that must evaluate to true for a particular action to be carried out. Applying this rule, we know the action will not be carried out if an individual only has a valid I.D. (TF), only has a boarding pass (FT), or has neither a valid I.D. nor boarding pass (FF). Table 1 displays truth-table assignments and real-world examples for each of the 16 fundamental Boolean logic operators.[1]

Much research on the modal concepts was conducted in the 1950s, 1960s, and 1970s, when the well-known definitional (e.g., logical rule) representational view for conceptual processing directed progress in the field (Bruner et al. 1956; Hull 1920; Murphy 2002). In testing the conceptual difficulty experienced when learning these operators, researchers construct a set of stimuli by systematically selecting two, three, or four visual (and occasionally auditory) stimulus dimensions. Then, they randomly select a modal concept and determine via truth-table analysis how each stimulus in the set is partitioned into one of two categories (see Table 1). Finally, they randomly present each stimulus to participants, instructing them to learn to which category each stimulus belongs upon receiving "correct" or "incorrect" feedback after each classification response. In general, learning difficulty for any particular modal concept may be operationalized in terms of the proportion of errors made, or number of trials needed, before participants

[1] For most of the examples provided in Table 1, note the arbitrary nature by which the particular assignment of dimensions and dimensional values were assigned to the truth-table structure. More specifically, we could have also processed the above instance of *conjunction* as, "Must have both a boarding pass and a valid I.D." Notice how simply switching the items alters the subsequent instances that do not satisfy this rule.

achieve a series of errorless classification responses, most often termed a “classification criterion” (Bourne 1970; Bourne and Guy 1968; Bruner et al. 1956).

One such investigation by Neisser and Weene (1962) assessed the learning difficulty for 10 of the 16 modal concepts—including *conjunction* (“and”), *inclusive disjunction* (“or”), *conditional* (“if, then”), and *biconditional* (“if and only if”). They hypothesized that the concepts were organized into a three-level hierarchy whereby concepts in lower levels of the hierarchy were easier to learn conceptually and acted as building blocks for concepts at higher levels in the hierarchy. To empirically test this hypothesis, they constructed four-character strings of consonants with the position of each character acting as a dimension with the possible character values of *J*, *Q*, *V*, *X*, and *Z*. They informed participants before each learning problem that two of the values (letters) would be important for solving the classification problem. In general, they found partial empirical support for the hypothesis that modal concepts belonging to higher levels in the proposed hierarchy would be more difficult to learn conceptually—operationalized as higher proportions of incorrect responses when classifying stimuli belonging to such concepts. However, they also found that concepts *within* a level (e.g., level two) might be differentially difficult to learn. Specifically, they discovered that, under this protocol, participants made more errors learning *conditional* than *inclusive disjunction* and more errors learning *inclusive disjunction* than *alternative denial*. These differences are particularly interesting considering that each of these three concepts are instances of the same structure type (see structure type $2_2[3]$ in Table 1): that is, the relationship between the dimensional values corresponding to each instance is the same (Feldman 2000, 2003; Goodwin and Johnson-Laird 2011; Vigo 2006, 2009, 2013, 2014). This idea of concept structure will be explained in greater detail shortly.

Similarly, Bourne (1970) investigated the learning difficulty of four primary modal concepts: *conjunction*, *inclusive disjunction*, *conditional*, and *biconditional*, and partially corroborated the hierarchy results of Neisser and Weene. Using a set of visual geometric stimuli, Bourne conducted two experiments following the “rule-learning” paradigm established by Haygood and Bourne (1965). Under this paradigm, participants are told the two relevant attributes for each learning problem and are tasked with discovering the logical rule connecting them via corrective feedback. Quantitatively, learning difficulty for any concept is operationalized as the number of trials until a preset learning criterion (i.e., 16 successive correct classifications) is reached. The results of both experiments conducted by Bourne (1970) reveal the following learning difficulty ordering (in terms of proportion of errors) for these four modal concepts: *conjunction* < *inclusive disjunction* < *conditional* < *biconditional*. This ordering has been partially corroborated using similar experimental paradigms and is an important contribution to how these modal concepts are studied in the categorization literature (Bourne and Guy 1968; Bourne et al. 1969; Bruner et al. 1956; Conant and Trabasso 1964; Dobson and Dobson 1981; Neisser and Weene 1962; Walls et al. 1975).

A few studies, however, have failed to replicate the Bourne (1970) ordering. For example, Reznick et al. (1978) utilized the rule-learning paradigm and tested the concept learning difficulty for the same set of logical rules as Bourne. Their main experimental manipulation involved presenting either spatially separated or spatially integrated stimuli to participants (the latter being analogous to stimuli utilized by Bourne and others). Regardless of whether the stimuli were spatially separated or integrated, they discovered the following learning difficulty ordering for the four rules: (*conjunction*, *inclusive disjunction*) < (*conditional*, *biconditional*). Not only does this ordering diverge from that discovered by Bourne, but *inclusive disjunction* and *conditional* are category instances belonging to the same structure type suggesting that the learning difficulty of these modal concepts, when viewed strictly from the standpoint of concept structure, may reveal yet another fundamental learning difficulty ordering. Indeed, many empirical studies have investigated the impact of concept structure on concept learnability (Feldman 2000; Goodwin and Johnson-Laird 2011; Nosofsky et al. 1994; Shepard et al. 1961; Vigo 2013). These researchers routinely operationalize categorization performance on structure types by averaging classification performance across their structure instances. Following this tradition, in the present study we wish to explore for the first time the learning difficulty of the 16 modal connectives from the standpoint of concept structure. Thus, we next describe in greater detail the distinction between structure instances and structure types.

The Boolean logic operators and structure types

In addition to being differentiable on their truth-value assignments, each of the Boolean operators in Table 1 may also be classified as an instance of a particular structure type (see leftmost column of Table 1; Feldman 2003; Higonnet and Grea 1958). This characterization is based on the inherent logical relationships between members that evaluate as positive examples of the operator. For example, an instance of *affirmation* may consist of either {**TT**, **TF**} or {**TT**, **FT**}, depending on which of the two properties (or dimensions) evaluates to true, and both are of the structure type $2_2[2]$-I.[2] Figure 1 displays the relationship

[2] The latter $D_n[p]$-Type notation introduced by Vigo (2013) indicates that the structure is defined over D dimensions (two in the case of the classical Boolean operators), that are *n*-ary (binary in the case of the classical Boolean operators), with *p* positive examples (two positive

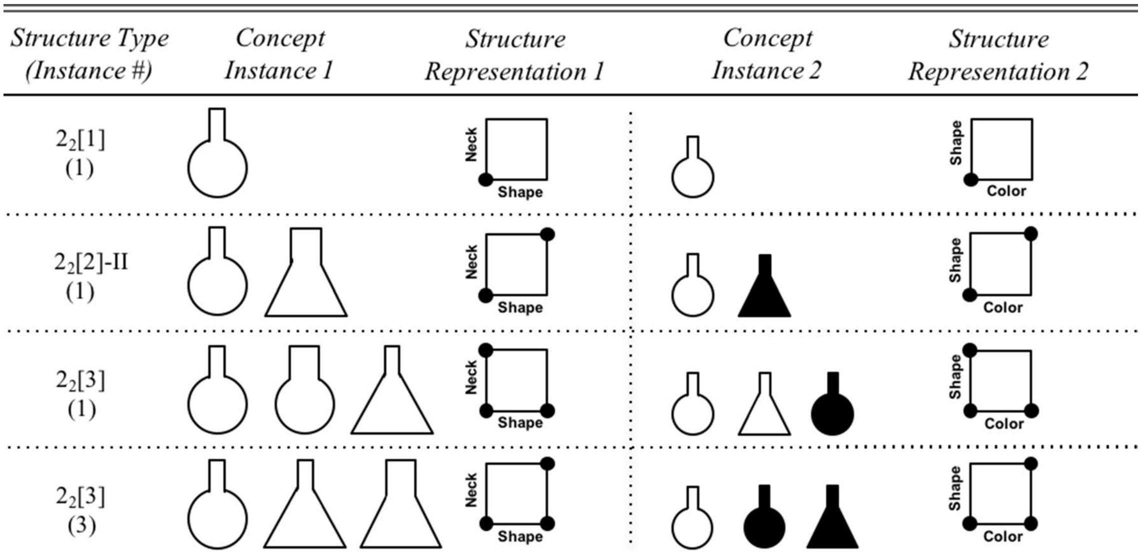


Fig. 1 Visual depiction of two sets of four structure instances. Columns "Concept Instance 1" and "Concept Instance 2" display a unique instance of each concept, where each concept is defined over the stimulus dimensions of *size* (small, large), *shape* (circular, triangular), *color* (black, white), and *neck width* (narrow, wide). Columns "Structure Representation 1" and "Structure Representation 2" display a Boolean square representation of each concept—highlighting the specific logical relationships objects share with each other in each concept instance. Note, that for "Concept Instance 1" and "Concept Instance 2," the dimensions of *size* and *color* and *size* and *neck width* are held constant and thus irrelevant in terms of classifying members from non-members, respectively. In the first column, "Instance #" refers to the instance order for each structure type as used in this paper and presented in Table 3

between structurally equivalent instances for two instances of structure type $2_2[3]$, and the column "logical structure" of Table 3 displays this relationship for all 16 modal concept instances. Note how this classification scheme results in several operators being instances of the same structure type. This equivalency is important as researchers routinely determine average proportion of classification errors for structure types upon averaging performance across their "structurally equivalent" instances (Bourne 1970; Feldman 2000; Nosofsky et al. 1994; Shepard et al. 1961; Vigo 2009, 2013, 2014).

A recent study by Kurtz et al. (2012), however, provides evidence that some instances belonging to the same three-dimensional structure types (e.g., "SHJ" or $3_2[4]$ category types; Shepard et al. 1961) may differ in their learnability. More specifically, Kurtz et al. (2012) observed differences in learnability depending on the nature of the stimulus dimensions and how they are systematically and logically assigned to the specific structure instances. For instance, some dimensions are easier to verbalize when forming logical rules for classifying members and non-members. If these dimensions are relevant for forming a valid logical rule, then classification will be easier for those particular structure instances. This logic may extend to other studies on the two-dimensional logic operators that report inconsistencies with respect to the learnability of their structurally equivalent instances (Neisser and Weene 1962; Reznick et al. 1978). We address this concern in the current research by utilizing stimulus dimensions that may be construed as equally verbalizable. Thus, we suggest the current work provides an unbiased and generalizable view into the learnability of the six structure types associated with the sixteen modal concepts.

The current investigation

As mentioned, one goal of the current research is to determine concept learnability from the standpoint of concept structure: that is, from the standpoint of the perceived relationships between the values of the dimensions that define the set of objects or exemplars (i.e., categorical stimuli) from which concepts are learned. Another impetus for the current work lies partly in the belief that the rule-learning protocol implemented by many researchers imposes limits on the interpretability and generalizability of subsequent empirical results (Bourne 1970; Dijkstra and Dekker 1982; Haygood and Bourne 1965; Reznick et al. 1978). First, researchers implementing the rule-learning paradigm notify participants, per trial, of the two relevant features (i.e., dimensional values) necessary to determine the rule for correctly

Footnote 2 (continued)

examples in the case of the affirmation operator above). The Roman numeral for Type is simply an arbitrary label to distinguish between logically distinct instances (e.g., structure types) belonging to the structure family.

partitioning the objects into the two categories. This facet of the rule-learning experimental design was created so different logical rules may be tested for their difficulty independent of processes related to attribute identification (Bourne and Haygood 1965). It is well known, however, that categorization performance is not solely reliant on attribute identification and rule-learning processes. Indeed, Bourne and Haygood (1965) demonstrated this fact with a complete-learning task where participants were not provided information regarding the relevant dimensions nor rule necessary to solve each learning problem.

This complete-learning task is most similar to the tasks employed in the current work where each participant is tasked with determining these relevancies themselves (Feldman 2000; Nosofsky 1984; Nosofsky et al. 1994; Rehder and Hoffman 2005). In the current experiments, we only informed participants of the four visual or auditory dimensions that made up each stimulus set. We required subjects to determine for themselves the relevant dimensions, attributes, and the rule that partitions the stimuli into the correct categories based on corrective feedback after every classification response. We posit that using four binary-valued stimulus dimensions and holding two constant for each learning problem should reduce the need to specify a priori the number of stimulus dimensions or the specific attributes per stimulus dimension per learning problem.

A second factor limiting the generalizability and interpretability of the results by Bourne (1970) and others is that often there was no time constraint per trial for each learning problem. This resulted in unequal amounts of time being spent across the rules. However, frequently we must make classification decisions quickly, thus arising the need to assess learning difficulty across the logic operators upon controlling for stimulus exposure. We followed this latter approach by fixing stimulus presentation across each of the 16 structure instances to three seconds for each classification trial (e.g., an individual classification response).

In addition to the two factors just discussed, there has been little experimentation to uncover whether the learnability of structure instances belonging to these structure types differs depending on the nature of the incoming sensory information (e.g., visual vs. auditory). Instead, research has primarily focused on varying stimulus information, such as the amount of relevant and irrelevant dimensions composing the stimuli (Archer 1962; Bulgarella and Archer 1962; Haygood 1965; Lordahl 1961; Pishkin and Shurley 1965). To clarify, a relevant dimension is necessary to successfully solve a classification problem, whereas an irrelevant dimension serves no role in solving such problems. For example, if we consider auditory stimuli defined over *timbre* (piano or violin) and *volume* (loud or soft) dimensions, one modal concept may consist of the category containing a loud piano tone (TT) and soft piano tone (TF; an instance of the operator *affirmation*). Here the *timbre* dimension is relevant to correctly classify the objects as it results in perfectly classifying the stimuli into two groups: all piano tones belong to one category, and all violin tones belong to another category. On the contrary, the *volume* dimension is irrelevant to successfully solving this problem and results in an average classification performance of 50% correct. In general, studies varying this type of stimulus information reveal a decrease in categorization performance (e.g., higher proportions of errors) when increasing the number of relevant and/or irrelevant acoustic dimensions (Bulgarella and Archer 1962; Pishkin and Shurley 1965).

A study by Haygood (1965), however, compared classification performance between visual and acoustic categories and revealed that increasing the number of relevant dimensions resulted in an increase in performance (e.g., lower proportions of errors) for visual stimuli, but not for acoustic stimuli. This may be explained using dual-coding theory (Paivio 1971, 1986), positing that incoming visual and verbal information is separated and processed differently, resulting in distinct mental representations. Or, as Haygood suggests, the result may be partly because people have considerable experience with visual coding, but little experience with auditory coding. Clearly, neither the result nor explanations negate the possibility that a single conceptual processing system processes both kinds of object stimuli. We address this possibility in the current study by conducting two categorization experiments on instances of the six structure types (for which the sixteen modal concepts serve as instances) with stimuli composed of either visual or acoustic dimensions. By comparing categorization performance across observers, we aim to test the hypothesis that concept formation, or the ability to generalize, is modality independent with respect to visual and acoustic information. In other words, that observers employ one conceptual system to process both acoustic and visual sets of objects.

In summary, the current work addresses several questions with respect to how observers learn visual and auditory modal concepts. First, is there a difference in learning difficulty between structure instances belonging to the same structure type? Second, does the Bourne ordering emerge with a modern categorization task that involves presenting to observers limited a priori information and a time constraint per classification trial? And finally, does the learnability of the 16 structure instances and six structure types depend on the nature of the incoming sensory (in our case, visual vs. auditory) information? Next, we briefly discuss two theories and formal models that may be used to predict categorization performance in the current tasks and provide tentative answers to these questions. Thereafter, we present the methodology and results of the current experiments.

Accounting for categorization behavior

A basic goal of categorization research is to account for how difficult any category (e.g., structure type) is to learn from a conceptual standpoint. Therefore, we utilized two successful theories and accompanying formal models of categorization behavior to characterize and account for the learning difficulties across the six structure types. First, the generalized context model (GCM; Nosofsky 1984, 1986) and associated theory assumes observers determine the similarity between a newly presented stimulus and category examples (i.e., exemplars) stored in memory by distributing attentional resources across the relevant dimensions in an optimal fashion—that is, in a manner consistent with minimizing classification errors. Specifically, the similarity between this new stimulus and the exemplars is computed in part using Shepard's (1987) exponential law of generalization combined with Luce's (1959) choice rule to determine the probability it will be classified correctly. Over the years, the model has successfully accounted for various results in traditional classification tasks, including the learning difficulty ordering of the widely studied "SHJ" structure types (Kruschke 1992; Nosofsky 1984; Nosofsky et al. 1994; Nosofsky and Johansen 2000; Nosofsky and Palmeri 1996; Shepard et al. 1961; Vigo 2013), and has even been applied to understand processes underlying social judgment (Halberstadt et al. 2011; Smith and Zarate 1992). Furthermore, the GCM remains the quintessential example of greatly parameterized process models that are based on similarity assessment and attention.[3]

In contrast, generalized invariance structure theory (GIST; Vigo 2013, 2014) posits that observers implicitly detect relational patterns in categorical stimuli, termed "categorical invariants," that subsequently inform rule-formation and concept learnability processing subsystems. The theory assumes these patterns are extracted via a process termed *dimensional binding,* whereby observers suppress (e.g., ignore) individual stimulus dimensions while engaging in partial similarity assessment between pairs of objects on the unsuppressed dimensions. Notice how this is a fundamentally different notion of similarity assessment than that posited to function in the GCM. The associated model (GISTM; Vigo 2013, 2014) posits that the degree of concept learning difficulty of a category is a function of both the amount of relational homogeneity or coherence (i.e., degree of categorical invariance) perceived in the category and the size (i.e., number of objects) of the category. As with the GCM, the model accounts for the ordering on the six "SHJ" categories and, unlike the GCM, for over 90% of the variance in the classification data across 84 structure types (Vigo 2013, 2014) without free parameters. For these and other reasons, Vigo (2013, 2014) has proposed the GISTM as a candidate law of conceptual behavior and is thus an ideal model to apply to the current study to predict and explain both visual and auditory categorization behavior (see "Appendix A" for more technical details of GIST and its associated formal model, the GISTM). Next, we present the methodology of the categorization task used to test the hypothesis that the same conceptual/generalizing cognitive subsystem is responsible for processing visual and acoustic categorical information.

[3] In contrast, more recent models of concept learning difficulty such as Feldman's algebraic complexity (Feldman 2006) and Vigo's QMV (Vigo 2006) are based on objective descriptions of the categorical stimulus and do not account well for recent empirical findings on concept learning difficulty (for further detail, see Vigo 2013, 2014).

Methods

Participants

Eighty-two participants were recruited from introductory psychology courses at Ohio University, 33 participants for the visual categorization task and 49 participants for the auditory categorization task. One auditory categorization participant reported having a hearing disorder and nine auditory categorization participants either withdrew from the study or did not finish the task in the time allotted. The remaining 39 participants in the auditory task reported having normal hearing and no history of hearing disorders. All participants were 18 years or older.

Stimuli

HP XW4600 workstations with Dell 1708FP 15-inch flat panel LCD monitors (5 ms. response time) were used to present sets of stimuli. Two types of stimuli were used, namely geometric flasks for the visual task and acoustic tones for the auditory task. The flasks varied over four binary dimensions: *size* (small or large), *shape* (circular or triangular), *color* (black or white), and *neck width* (narrow or wide). Visual examples of these flasks are shown in Fig. 1, in Vigo and Basawaraj (2013), and in Vigo (2013, 2014). The tones varied over four binary dimensions: *duration* (500 or 1000 ms), *pitch* (C2 or C3), *timbre* (piano or violin), and *loudness* (soft [~ 60 dB] or loud [~ 70 dB]). Importantly, empirical results from Miskiewicz and Rakowski (2012) indicate that the chosen pitch levels are easily differentiated. To approximate the levels of loudness, a mobile speaker was placed roughly 2.5 cm from the headphone speaker, simulating the approximate distance between the headphone speaker and the eardrum (Goode 2001). The Decibel 10th mobile application provided approximate measurements for each level of loudness.

Table 2 Six sets of visual (flask) and auditory (tone) stimuli presented to participants

Stimulus set	Visual		Auditory	
	Dimension 1	Dimension 2	Dimension 1	Dimension 2
1	*Shape* (Circular, Triangular) All: [Large, White]	*Neck width* (Wide, Narrow)	*Loudness* (~ 60 dB, ~ 70 dB) All: [C3, Short]	*Timbre* (Piano, Violin)
2	*Neck width* (Wide, Narrow) All: [Small, Circle]	*Color* (Black, White)	*Timbre* (Violin, Piano) All: [C3, ~ 60 dB]	*Duration* (1000 ms, 500 ms)
3	*Size* (Large, Small) All: [Black, Wide]	*Shape* (Triangular, Circular)	*Loudness* (~ 70 dB, ~ 60 dB) All: [Violin, 1000 ms]	*Pitch* (C3, C4)
4	*Color* (White, Black) All: [Triangular, Narrow]	*Size* (Large, Small)	*Duration* (500 ms, 1000 ms) All: [Piano, ~ 70 dB]	*Pitch* (C3, C4)
5	*Neck Width* (Wide, Narrow) All: [Black, Triangular]	*Size* (Small, Large)	*Pitch* (C4, C3) All: [~ 70 dB, 1000 ms]	*Timbre* (Violin, Piano)
6	*Color* (White, Black) All: [Small, Narrow]	*Shape* (Circular, Triangular)	*Loudness* (~ 60 dB, ~ 70 dB) All: [Piano, C3]	*Duration* (500 ms, 1000 ms)

We created each set of stimuli upon combining values across dimensions. For instance, Visual set 1 would consist of the following four stimuli: {Circle and Wide, Circle and Narrow, Triangle and Wide, Triangle and Narrow}. Because we had to control for values on two irrelevant dimensions, each of these four stimuli would also be large and white (an admittedly arbitrary choice for these two-dimensional values)

For any one trial only two dimensions varied, while the two remaining dimensions were held constant; this procedure results in sets of four stimuli for constructing different instances of each of the six structure types. Each set was generated based on the 16 truth-table assignments shown in Table 1, and Table 2 lists the six bi-dimensional combinations of the visual and acoustic sets randomly sampled and displayed to participants.[4] And despite there being six ways to choose two dimensions out of four (e.g., *size* and *shape*, *tone* and *timbre*), each participant was tested on a random four of the six possible combinations provided in Table 2 for each of the 16 structure instances. Thus, participant 1 may receive stimulus sets 1, 3, 4, 6 for instance 1 and stimulus sets 2, 3, 4, 5 for instance 2, whereas participant 2 may receive stimulus sets 2, 3, 5, 6 for instance 1 and stimulus sets 2, 3, 4, 6 for instance 2. Finally, each participant categorized sets belonging to 64 structure instances (16 unique instances × 4 sets each) for either the visual or acoustic stimuli, but not both. Each stimulus was presented individually, with a neutral gray background displayed throughout the entire experiment. Participants in the auditory categorization task used KOSS SB45 headphones to listen to the auditory tones.

Procedure

Upon giving consent, the researcher explained to the participant(s) that they were to learn the preferences of either an art collector (flasks) or a musician (tones). They were told that they would be presented with individual figures (sounds) that varied by size, shape, color, and neck width (tone, timbre, duration, and loudness). Upon the presentation of each figure (sound), they were instructed to press the LEFT mouse button if they thought the figure (sound) was liked by the art collector (musician), or the RIGHT mouse button if they thought the figure (sound) was disliked. Additionally, they were told that initially they would not know the preferences of the art collector (musician) but would learn through corrective feedback after each response. After answering any questions, the participants were directed to the computers to complete the task.

[4] When choosing two relevant attributes from the eight total attributes constituting the four-dimensional stimuli, there are 48 combinations (8 attributes × 6 attributes, since the second attribute must be from a different dimension). For example, one of the 48 combinations involves the first relevant attribute of *size* = small and the second relevant attribute of *color* = white. A second, albeit similar, combination reverses the order, resulting in *color* = white for the first relevant attribute and s*ize* = small for the second relevant attribute. In the current experiments, we randomly selected 6 of these 48 combinations for constructing the visual and acoustic stimulus sets in the current study.

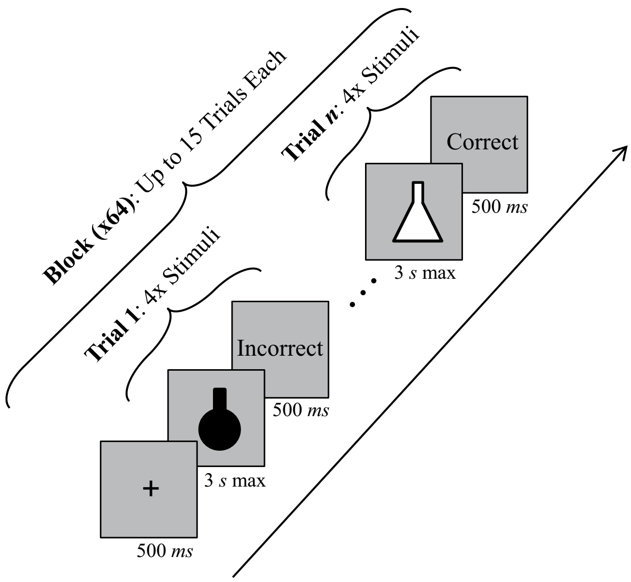


Fig. 2 Shown is a visual depiction of one block of categorization trials. For each trial within a block, participants made a series of four classification decisions (one for each of the four unique stimulus objects (flasks or tones) assigned to that specific block). A block of trials was considered "learned" if a participant completed four consecutive trials of errorless classification decisions (i.e., 16 decisions). Participants completed a maximum of 15 trials (60 decisions) for a block if they failed to reach the four trial learning criterion

Once seated, participants first completed a training session, which was conducted to ensure that participants understood the task. The training session consisted of presenting four blocks of trials, which were sampled at random from the 64 structure instances tested in the experimental block. Within each block, one trial involves the presentation of the four figures (sounds) presented one at a time (see Fig. 2). Upon the presentation of each figure (sound), participants had three seconds to make a response, and no response was counted as an incorrect response. Participants received corrective feedback after every response and had to reach a classification criterion of 16 consecutively correct classification responses (i.e., four perfect trials) to continue to the next block of trials. If the criterion was not met after 60 classification decisions (i.e., 15 trials), the experiment continued onto the next block of trials. Between blocks, participants viewed a neutral gray screen for three seconds, thus acting as a reset before the next block of trials.

Once completed with the four block training session, participants were asked whether they had any questions and subsequently began the experimental session. The experimental session consisted of 64 blocks of trials (16 structure instances × 4 combinations of dimensions), randomly sampled for each participant. After completion of the experimental session, participants were debriefed and thanked for their time.

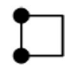
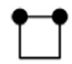
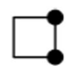
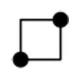
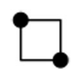
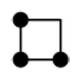
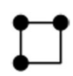
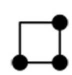
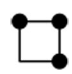
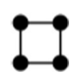
Results

Analyses of the four modal concepts studied by Bourne (1970)

In calculating performance for each of the 16 structure instances, we averaged categorization performance across the four combinations of dimensions per subject per structure instance ($N = 33$ subjects × 4 combinations = 132 blocks for Flask stimuli; $N = 39$ subjects × 4 combinations = 156 blocks for Tone stimuli). The descriptive statistics for each structure instance are provided in Table 3. Regarding these 16 structure instances, we only report statistical analyses regarding the data obtained for the four structure instances studied by Bourne (1970) and many others. This is because our experimental design, as it was currently implemented, does not permit separate pairwise statistical analyses across structurally equivalent instances. This is mainly due to the arbitrary nature by which stimulus dimensions and attributes provided in Table 2 were mapped onto the truth-table assignments provided in Table 1. Despite this fact, this only proves consequential for the four concepts studied by Bourne (1970) when comparing performance between the instances of *inclusive disjunction* and *conditional*.

We conducted Bayesian paired-samples t tests using JASP (JASP Team 2016) upon collapsing across stimuli (see nonsignificant interaction in next section) to determine statistical differences among the four modal concepts originally studied by Bourne (1970). Supplemental analyses associated with these statistical tests are shown in Fig. 6 of "Appendix B". We obtained the following learning difficulty ordering for the four structure instances: (*conjunction*, *inclusive disjunction*, *conditional*) < *biconditional*. In other words, there is not strong evidence to conclude that categorization performance differed between *conjunction* (instance 2), *inclusive disjunction* (instance 12), and *conditional* (instance 14), least favorable Bayesian paired-samples t test: $BF_{10} = 0.417$, 95% CI for δ: [− 0.05, 0.41]. However, we did find very strong evidence that *Biconditional* (instance 10) was harder to learn than the other three concepts, least favorable Bayesian paired-samples t test: $BF_{10} = 2.36 \times 10^9$, 95% CI for δ: [0.67, 1.25]. Note that this ordering does not replicate the ordering found by Bourne (1970) and other researchers, namely: *conjunction* < *inclusive disjunction* < *conditional* < *biconditional*. We attribute the difference in the orderings to the methodological differences, which were discussed at length in Introduction, between the current study and prior studies on these four modal concepts. Additionally,

Table 3 Average number of trials to solution for all 16 structure instances and six structure types

Structure type	Structure instance [#]	Logical structure	Average trials to criterion (SD)			
			Auditory	Visual	Combined	Type
2_2[0]	[None] [1]		1.26 (0.47)	1.06 (0.28)	1.17 (0.41)	1.17 (0.41)
2_2[1]	00 [2]		4.28 (2.06)	3.02 (1.79)	3.65 (2.03)	3.40 (1.79)
	01 [3]		3.02 (1.56)	3.48 (1.55)	3.25 (1.56)	
	10 [4]		3.41 (1.82)	3.47 (1.49)	3.44 (1.66)	
	11 [5]		3.03 (1.63)	3.47 (1.49)	3.25 (1.88)	
2_2[2]-I	00, 01 [6]		3.35 (2.28)	3.56 (1.81)	3.45 (2.07)	3.10 (1.91)
	00, 10 [7]		3.15 (2.23)	3.04 (1.94)	3.09 (2.09)	
	01, 11 [8]		2.36 (1.10)	3.05 (1.49)	2.71 (1.33)	
	10, 11 [9]		3.57 (2.22)	2.79 (1.65)	3.18 (2.01)	
2_2[2]-II	00, 11 [10]		6.96 (3.68)	7.00 (3.10)	6.98 (3.40)	7.09 (3.29)
	01, 10 [11]		7.16 (2.98)	7.24 (3.49)	7.20 (3.20)	
2_2[3]	00, 01, 10 [12]		4.22 (2.42)	3.75 (2.07)	3.99 (2.27)	3.70 (2.09)
	00, 01, 11 [13]		3.73 (2.40)	3.55 (2.35)	3.64 (2.36)	
	00, 10, 11 [14]		3.92 (2.24)	3.27 (1.61)	3.60 (1.99)	
	01, 10, 11 [15]		3.55 (1.75)	3.54 (1.68)	3.54 (1.70)	
2_2[4]	00, 01, 10, 11 [16]		1.03 (0.76)	0.64 (0.33)	0.85 (0.63)	0.85 (0.63)

The "0" and "1" values in the "Structure Instance" column represent the stimulus value for dimension 1 and dimension 2, respectively. For instance, instance 6 consists of two stimuli: a stimulus that has value "0" on dimension 1 and value "0" on dimension 2, and a stimulus that has value "0" on dimension 1 and value "1" on dimension 2. Also note that the remaining stimuli ("10" and "11") are non-members of this particular structure instance

the equality in classification performance observed with respect to instances of *inclusive disjunction* and *conditional* should be interpreted with caution due to the arbitrary nature by which stimulus dimensions and stimulus attributes were assignable to these instances. Thus, based on the current implementation of our experimental design, it is not unreasonable to suggest that participants treated both types of structure instances similarly. The remaining statistical differences were mainly the result of differences between structure types, which we report next.

Analyses of the six structure types

As stated previously, researchers routinely determine classification performance for structure types upon averaging performance across structurally equivalent instances (Shepard et al. 1961; Feldman 2000; Nosofsky et al. 1994; Vigo 2013, 2014). Earlier, we questioned the legitimacy of this approach and cited empirical evidence reporting different categorization performance on structurally equivalent instances (Kurtz et al. 2012; Neisser and Weene 1962; Reznick et al. 1978). Currently, we observed only one statistical difference between structure instances belonging to the same structure type (out of the 19 within-type pairwise comparisons). This difference was between instance 6 and instance 8 of structure type 2_2[2]-I, $BF_{10} = 21.52$, 95% CI for δ: [0.15, 0.62]. Because this was the only comparison yielding a Bayes Factor greater than 10 (and in actuality, greater than 2), the current results provide support for the procedure of averaging across structurally equivalent instances (at least with respect to the 16 two-dimensional instances tested in the current study). The average trials to criterion for each of the six structure types are shown in the rightmost column of Table 3, and supplemental analyses for the statistical tests reported in this section are shown in Fig. 7 of "Appendix B." Next, we report statistical analyses regarding whether the nature of the incoming stimulus information (e.g., visual vs. auditory) affects categorization performance across these six structure types.

A 2 (stimulus) × 6 (structure type) repeated-measures Bayesian ANOVA was conducted with stimulus (visual, acoustic) as the between-subjects factor and structure type as the within-subjects factor. The model that best accounted for the data included only the main effect term for structure type, BF = 16.05, p (M ∣ Data) = .80, error % = 0.224. In other words, we observed strong evidence of differences in categorization performance across the six structure types; however, these differences did not depend on whether the categorical stimuli consisted of visually defined flasks or acoustically defined tones. Additionally, the correlation between the visual and auditory modalities with respect to classification performance across the six structure types was approximately 1, $t(4) = 33.68$, $p < .001$ [95% CI: .983, 1]. Figure 3 displays this relationship across both modalities for both the 16 structure instance (top panel) and the six structure types (bottom panel).

Additional pairwise Bayesian t tests on the six structure types reveal strong evidence that structure type 2_2[4], consisting of all objects (instance 16), required less trials until the criterion was reached compared to structure type 2_2[0], consisting of no objects (instance 1), $BF_{10} = 67.13$, 95% CI for δ: [0.19, 0.67]. The Bayes factor (BF_{10}) reveals there is approximately 67 times more evidence favoring the alternative hypothesis (difference between the two structure types) than evidence favoring the null hypothesis. Second, both of these structure types required fewer trials until the criterion was reached when compared with each of the remaining four structure types, median $BF_{10} = 2.64 \times 10^{20}$, median 95% CI for δ: [− 2.08, − 1.35].

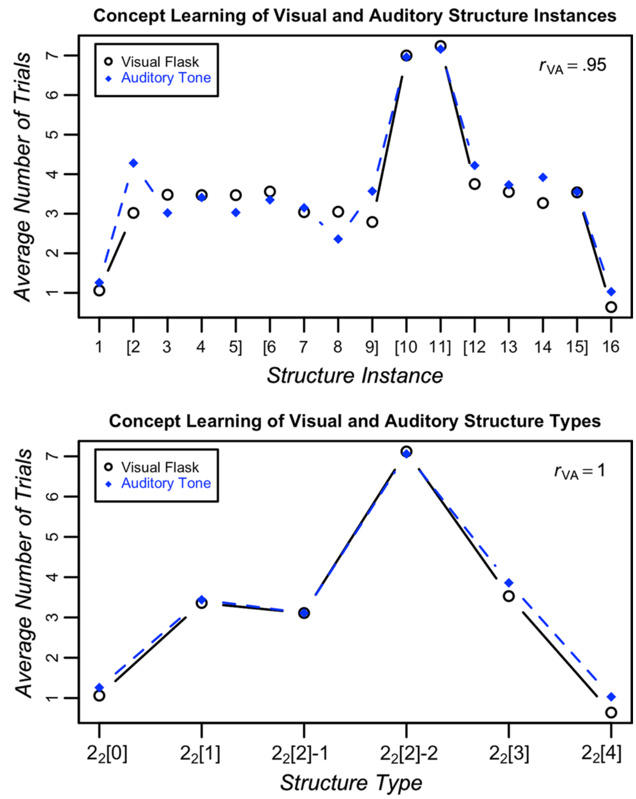


Fig. 3 Shown in the top row is a comparison of the learning difficulty of the 16 structure instances for both the visual (flask) and auditory (tone) stimuli. The correlation between performance on the flask and tone stimuli is .95, indicating a strong relationship in performance across the 16 structure instances between on the two types of stimuli. Shown in the bottom row is a comparison of the learning difficulty of the 16 structure instances upon being collapsed into their respective structure type for the flask and tone stimuli. The correlation between performance on the flask and tone stimuli upon collapsing across structure type is approximately 1

Regarding the four non-trivial structure types, we observed significantly more trials to criterion for instances belonging to type 2_2[2]-II than for the remaining three structure types, least favorable Bayesian independent samples t test: $BF_{10} = 2 \times 10^{29}$, 95% CI for δ: [1.09, 1.53]. This is not surprising given that the two structure instances belonging to type 2_2[2]-II abide by logical truth-table assignments for *biconditional* (instance 10) and its negation, *exclusive disjunction* (instance 11). As stated previously, we did not obtain sufficient evidence to conclude that these two structurally equivalent instances differed in learnability, $BF_{10} = 0.149$, 95% CI for δ: [− 0.28, 0.16]. Further, we found strong evidence that performance on type 2_2[2]-I required fewer trials to criterion than instances belonging to type 2_2[3], $BF_{10} = 44.83$, 95% CI for δ: [− 0.45, − 0.13]. Finally, learning difficulty for type 2_2[1] did not differ significantly from learning difficulties associated with types 2_2[2]-I or 2_2[3], least favorable Bayesian independent

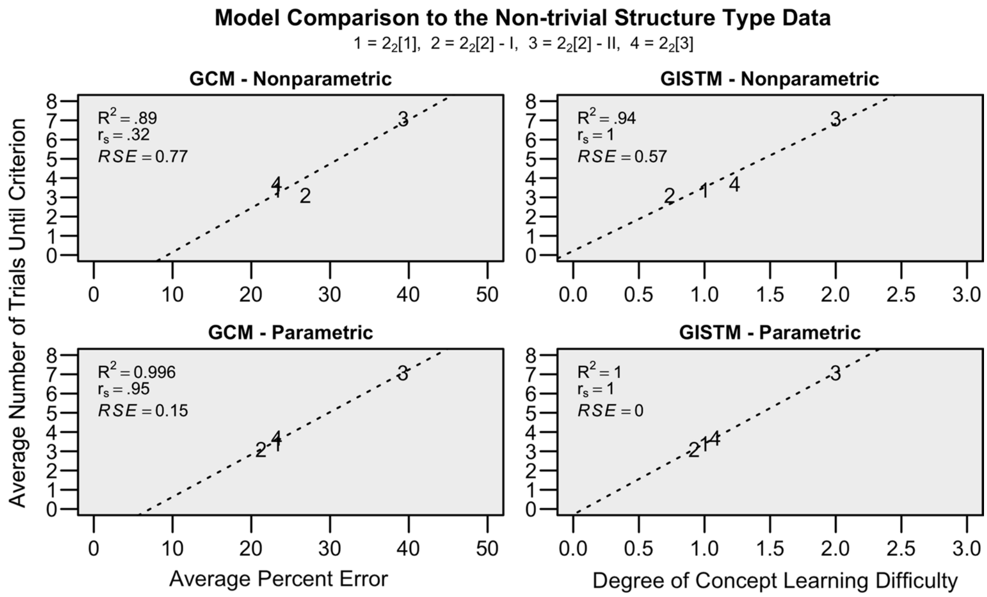


Fig. 4 Plotted are linear regression fits for each model across the four non-trivial structure types (i.e., $2_2[1]$, $2_2[2]$-I, $2_2[2]$-II, and $2_2[3]$). Nonparametric model fits for both the GCM (Nosofsky 1984, 1986) and the GISTM (Vigo 2013, 2014) are shown in the first row, whereas parametric model fits for both models are shown in the second row

samples t test: $BF_{10} = 0.527$, 95% CI for δ: [− 0.01, 0.32]. We obtained the following ordinal relationship between the six structure types with respect to average amount of trials needed to reach the classification criterion: $2_2[4]$ (0.85) < $2_2[0]$ (1.17) < $2_2[2]$-I (3.10) < $2_2[1]$ (3.40) < $2_2[3]$ (3.70) < $2_2[2]$-II (7.09). Next, we present a comparison of how two different categorization models account for these results across the four non-trivial structure types.

Model comparison

A final analysis in the current investigation is the associated fit between the observed categorization performance and predictions made by the GCM with and without free parameters (GCM; GCM-NP; Nosofsky 1984, 1986) and the GISTM with and without free parameters (GISTM; GISTM-NP; Vigo 2013, 2014) on the four non-trivial structure types. The variant of the GCM without optimized attention weights may be interpreted such that the distribution of attentional resources is uniform across dimensions (see Nosofsky 1984, 1986 for details). This assumption is consistent with the fact that attentional biases are canceled out in our experimental design by the random sampling of structure instances.

As shown in Fig. 4, both the GCM and GISTM provide an excellent account of the structure type data, regardless of whether the underlying parameters of each model are optimized. Despite the generalizability of the results being limited (because we are comparing performance on only four structure types), the two variants of the GISTM slightly outperform the two variants of the GCM. First, the non-parameterized GISTM accounts for approximately 94% of the variance in learnability among the four types, $t(2) = 5.51$, $p = .031$, whereas the non-parameterized GCM accounts for approximately 89% of the variance in learnability, $t(2) = 3.93$, $p = .059$. Spearman rank-order correlations favor the GISTM-NP ($r_s = 1$) over the GCM-NP ($r_s = .32$), and it is worth noting that the GCM without optimized attention weights cannot account for type 2[2]-I being significantly easier to learn than type 2[3].

Upon estimating the lone scaling parameter (k) and the two sensitivity weights (α values) for the GISTM (provided in Table 4), we find the parameterized GISTM accounts for approximately 100% of the variance in learnability among the four structure types, while the spearman rank-order correlation remains at 1. The parameterized GCM consisting of the lone estimated scaling parameter (c) and four estimated attention weights (α values) accounts for approximately 99.6% of the variance in learnability, $t(2) = 20.98$, $p < .001$, resulting in a much-improved spearman rank-order correlation of approximately .95. As can be seen in Table 4, the aforementioned assumption regarding attentional biases being canceled out as a result of the random sampling of structure instances is largely supported. Indeed, type 2[2]-I resulted in the only set of optimized attention weights that were not approximately .5 ($\alpha_1 = .654$, $\alpha_2 = .346$). In conclusion, upon considering the model complexity trade-off between number of estimated parameters and predictive capability without estimated parameters, we believe that there is a significant advantage favoring the GISTM in accounting for the present results. Furthermore, we believe that this advantage in performance lends some support to one of the key hypotheses of GIST that the detection of

Table 4 Optimized parameter values for the GCM (Nosofsky 1984, 1986) and GISTM (Vigo 2013, 2014) on the four structure types

	GCM			GISTM		
Structure type	Scaling (c)	Attention (α_1)	Attention (α_2)	Scaling (k)	Sensitivity (α_1)	Sensitivity (α_2)
2[1]	2.00	.50	.50	4.05	.61	.44
2[2]-I		.654	.346			
2[2]-II		.50	.50			
2[3]		.50	.50			

One scaling parameter was estimated for each model across the four structure types. For the GCM, four attention weights were estimated upon constraining the two attention weights per structure type to sum to 1. For the GISTM, two sensitivity weights were estimated (one per each dimension) and each was constrained to lie between [0, 1]

invariance patterns in categorical stimuli (i.e., categories of objects or exemplars) underlies conceptual behavior.

Discussion

A central hypothesis of the current study was that a single system is responsible for processing categorical information associated with the six structure types defined by two binary-valued stimulus dimensions, regardless if that information is presented visually or acoustically. Initial evidence for this hypothesis comes from the nonsignificant interaction between the two factors of *stimulus* and *structure type*, thus indicating that the same ordering between the 6 structure types was obtained for both types of stimulus sets. This is further supported by the strong positive correlations of .95 and 1 between the categorization performances of the structure instances and upon being collapsed into their structurally equivalent types, respectively, for the visually presented stimuli compared to the acoustically presented stimuli (see Fig. 3). Additionally, the learning difficulty ordering for the four modal concepts studied by Bourne (1970) and others—*conjunction*, *disjunction*, *conditional*, and *biconditional*—was also found to be identical regardless of which sensory modality was receiving the information [(*conditional*, *conjunction*, *disjunction*) < *biconditional*)]. In total, the strong positive relationship, the identical learning difficulty orderings, and the accurate predictions of the GISTM and GISTM-NP all provide converging evidence in support of the hypothesis that the six structure types are processed with a single unified conceptual system able to generalize visual and acoustical information in a fundamentally equivalent manner. Accordingly, a primary goal of this system may be to detect the invariance-based relational patterns (i.e., categorical invariants) inherent to sets of object stimuli that facilitate concept learnability and may subsequently be used to inform language processes related to rule-formation (Vigo 2013, 2014).

Another goal of the current work was to assess the validity of averaging categorization performance across structurally equivalent category instances. In the current study, we restricted our attention to structure instances belonging to the four non-trivial structure types defined by two binary-valued stimulus dimensions. We found strong empirical support for this procedure, with only one difference (out of a possible 19) emerging between instances belonging to the same structure type.

Notwithstanding, a future research direction should be to investigate the robustness and generality of the current results by exploring categorization behavior of categories of increased complexity (e.g., three- or four-dimensional). Additionally, a *parainformative* categorization task (Feldman 2000; Vigo 2014) could be conducted to reveal the nature of how the relational patterns in acoustic stimuli are extracted when all of the information is available simultaneously, rather than presented sequentially. One such investigation presented at a conference for the *Acoustical Society of America* by Vigo et al. (2012) addressed the visual and acoustic concept learning difficulty of three-dimensional categories consisting of four objects (Shepard et al. 1961) upon adopting the parainformative experimental design. Specifically, they reported a general increase in categorization errors for the auditory concepts when compared to their visual counterparts. They also discovered a small discrepancy among the visual and acoustic learning difficulty orderings of the six tested structure types. GIST accounts for these slight differences in learning by positing specific interactions between working memory capacity and categorical invariance pattern extraction processing, thus providing additional support for the hypothesis of a single unified conceptual system for processing acoustic and visual information. In conclusion, we acknowledge there is much to learn regarding visual and acoustic conceptual processing of categorical stimuli. We hope the current evidence in favor of a universal conceptual system motivates more research into the nature of modal concepts across the various sensory modalities and into their relative learning difficulty.

Appendix A

The generalized invariance structure theory model (GISTM)

Generalized invariance structure theory, or GIST (Vigo 2013, 2014), proposes that observers are invariance pattern detectors. In other words, observers detect abstract symmetries inherent in the dimensional structure of a category of objects with the ultimate aim of efficiently determining the degree of diagnosticity of each of the category's relevant dimensions. The observer is then able to ascertain or assess which dimensions should be used in the formation of concept learning rules. As such, the ability to detect invariance patterns in categorical stimuli is a necessary precursor to concept formation in GIST. The core model of the theory is referred to as the "generalized invariance structure theory model," or GISTM. The parameterized variant of the model we are employing (see Vigo 2014 and supplementary materials to Vigo 2013) is expressed as follows:

$$\psi(\mathrm{X}) = pe^{-k\widehat{\Phi}_{\alpha}^{2}(\mathrm{X})} \qquad (1)$$

where ψ is the degree of perceived learning difficulty of a continuous or dichotomous category X, p is the cardinality or size of the categorical stimulus, D is the number of dimensions used to define X, and $\widehat{\Phi}_{\alpha}$ is the degree of perceived categorical invariance determined by the proportion of categorical invariants $H_{[d]}$ (X) in X with respect to dimension d ($1 \leq d \leq D$) as follows.

$$\widehat{\Phi}_{\alpha}(\mathrm{X}) = \left[\sum_{d=1}^{D}\left[\alpha_d H_{[d]}(\mathrm{X})\right]^2\right]^{1/2} \qquad (2)$$

Note that this parameterized version of the GISTM includes a discrimination parameter k ($k \geq 0$) and an invariance detection sensitivity parameter α_d per dimension d (where for any d, $0 \leq \alpha_d \leq 1$). On the other hand, the nonparametric variant of the model (i.e., the GISTM-NP) does not feature any free parameters and takes the following forms (where $D_0 = 2$ and $\frac{D_0}{D}$ is a category structure discrimination index determined relative to the smallest number of dimensions of a category structure: namely, two):

$$\psi(\mathrm{X}) = pe^{-\frac{D_0}{D}\widehat{\Phi}^{2}(\mathrm{X})} \qquad (3)$$

$$\widehat{\Phi}(\mathrm{X}) = \left[\sum_{d=1}^{D}\left[H_{[d]}(\mathrm{X})\right]^2\right]^{1/2} \qquad (4)$$

Essentially, the invariance detection sensitivity parameter reflects the effectiveness of a lower level cognitive mechanism of invariance detection referred to as "dimensional binding." Dimensional binding requires that similarity assessment be relativized by the process of systematically and completely suppressing each relevant dimension during similarity comparisons. For a formal specification of this mechanism, please refer to Vigo (2013, 2014).

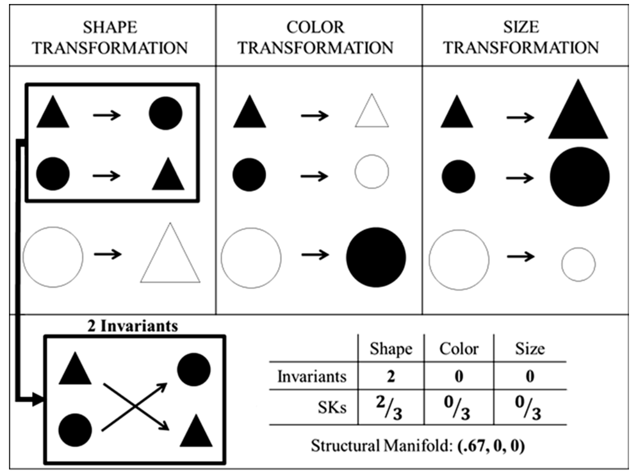


Fig. 5 Summary of the process of detecting categorical invariants using a simple category structure (*small black triangle*; *small black circle*; *large white circle*) consisting of three objects and three dimensions

Figure 5 shows the process of detecting invariants using a simple category structure (*small black triangle*; *small black circle*; *large white circle*) consisting of three objects and three binary dimensions. In the original structural account of the model (Vigo 2009), a differential operator generates the degree of partial invariance by perturbing dimensions of categorical stimuli. These perturbations are dimensional transformations that determine the number of invariants per dimension. The number of invariants per dimension equals the number of common objects between the original and perturbed categories. Thus, upon the shape transformation in Fig. 5 we see that the *small black circle* and the *small black triangle* remain after the perturbation. Upon the color and size transformations, however, no objects are common to the original and perturbed sets.

This differential operator is interpreted as a cognitive mechanism or cognitive operator $H_{[d]}$(X) via the process of dimensional binding mentioned above in Vigo's (2013) generalization of the model. The invariance detection operator generates one structural kernel (SK) per dimension where SKs are the proportion of invariant objects to the total number of objects in the category. The structural manifold of the category is found by computing the proportion of categorical invariants (with respect to each dimension) to the total number of objects and arranging these proportions as a vector.

In general, this process determines how relatively essential a given dimension is in terms of characterizing category membership. Simply, objects either remain or are eliminated after a perturbation. Dimensions with a relatively greater number of eliminated objects after perturbation are more essential for determining category membership. Alternatively, dimensions with a relatively greater number of objects that remain after perturbation are relatively non-essential for determining category membership. Therefore, the structural manifold obtained in our Fig. 5 example indicates that color and size are essential for classification, whereas shape is relatively non-essential. At this stage, the observer has the information that is necessary for forming an efficient classification rule.

To determine the degree of learning difficulty of the category in Fig. 5, we use the GISTM-NP (Eqs. 3 and 4) as follows. First, using Eq. 4, we compute the global degree of categorical invariance $\widehat{\Phi}$ using the structural manifold (.67, 0, 0) of the category (we shall refer to the category as X). Recall that $H_{[d=1]}$ (X) = 2/3 ≈ .67, $H_{[d=2]}$ (X) = 0, and $H_{[d=3]}$ (X) = 0. We then get the following:

$$\widehat{\Phi}(\mathrm{X}) = \left[\sum_{d=1}^{D} \left[H_{[d]}(\mathrm{X})\right]^2 \right]^{1/2} = [[2/3]^2 + [0]^2 + [0]^2]^{1/2} \approx .67 \qquad (5)$$

We can now compute the degree of learning difficulty ψ of X using Eq. 3 above and get:

$$\psi(\mathrm{X}) = pe^{-\frac{D_0}{D}\widehat{\Phi}^2(\mathrm{X})} = 3e^{-\frac{2}{3}(.67)^2} \approx 2.22 \qquad (6)$$

Appendix B

See Figs. 6 and 7.

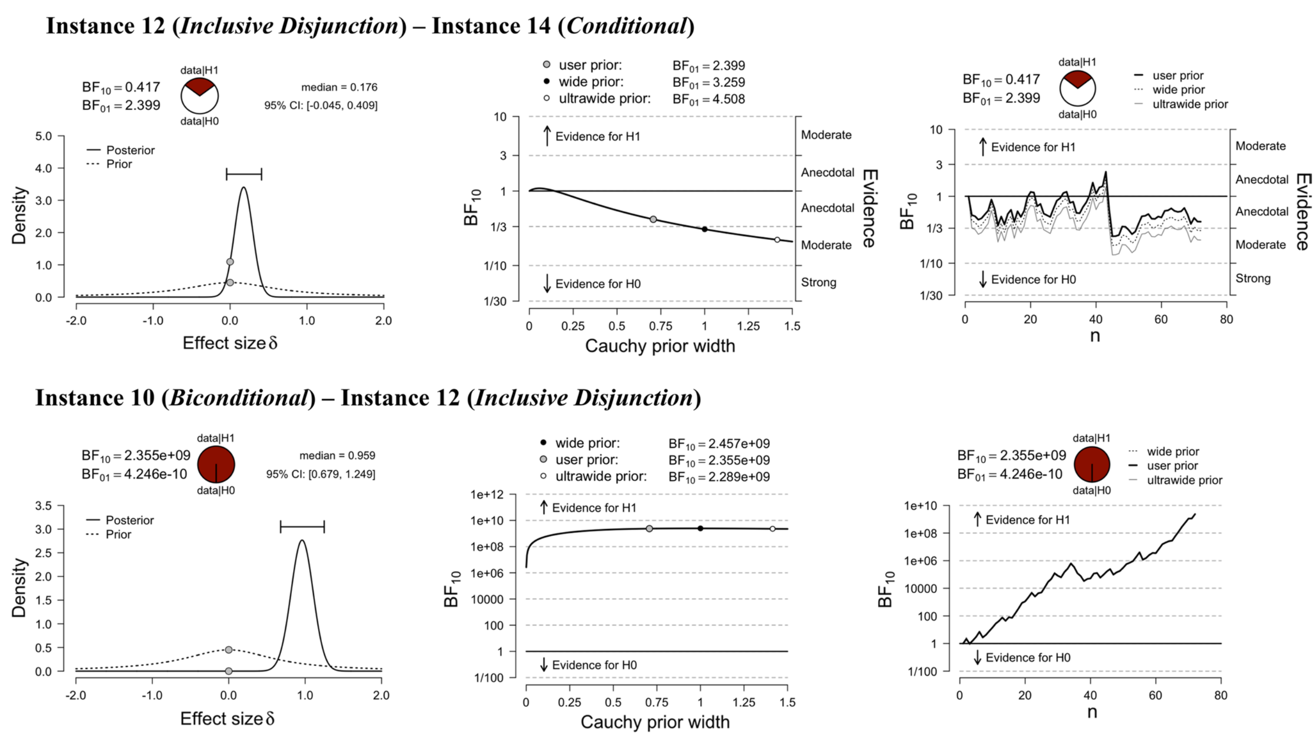


Fig. 6 Plotted are supplemental analyses for the Bayesian paired *t* test results reported in the "Analyses of the 16 structure type instances" in "Results" section. The top row displays plots for the *t* test result that was closest to providing evidence against the null hypothesis of no statistical difference between three structure type instances of *conjunction*, *inclusive disjunction*, and *conditional*. The bottom row displays plots for the *t* test result most favorable to the null hypothesis of no statistical difference between each of these three structure instances and *biconditional*. The plots were created using JASP (JASP Team 2016)

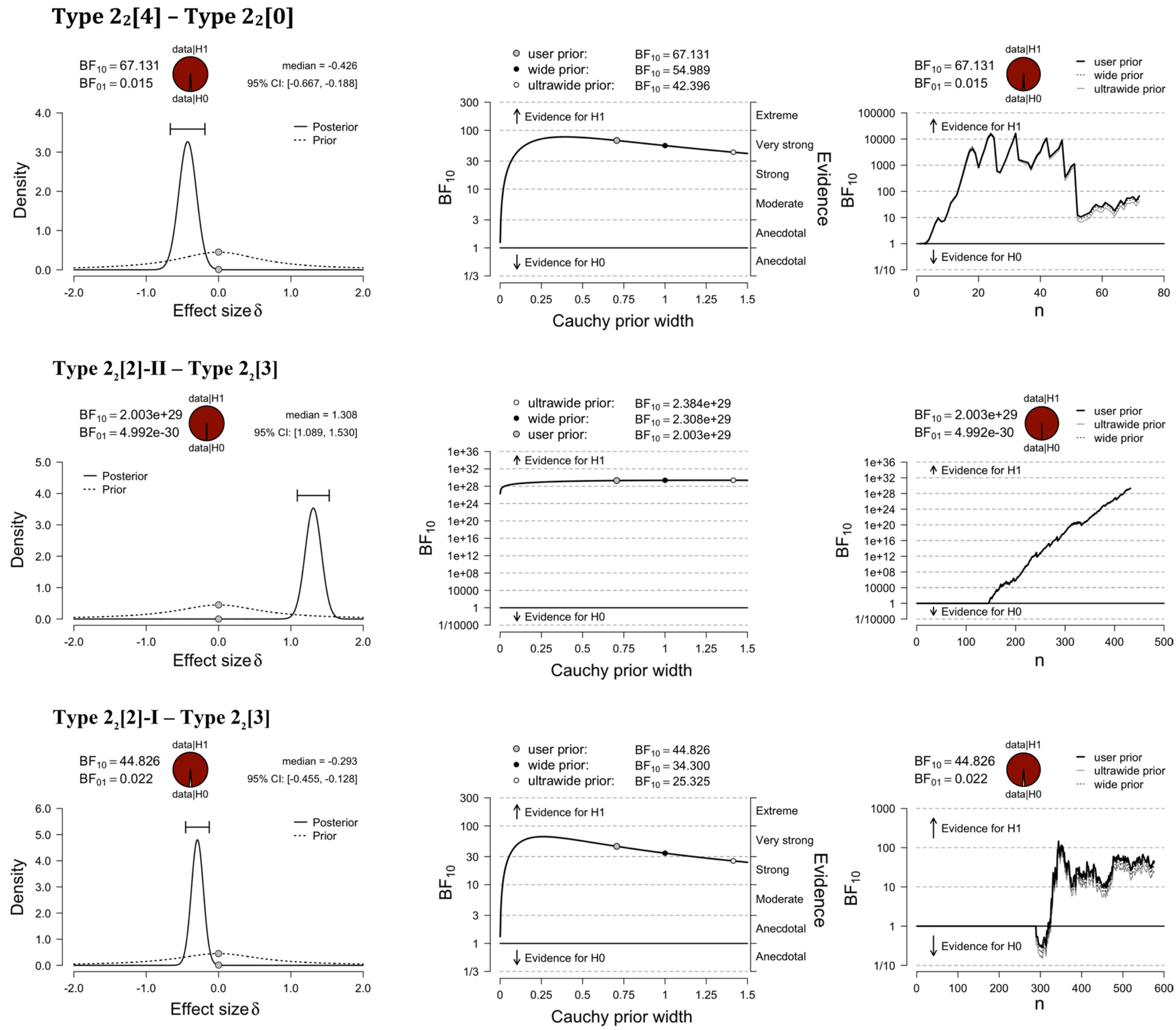


Fig. 7 Plotted are supplemental analyses for the Bayesian independent samples t test results reported in the “Analyses of the six structure types” in “Results” section. The plots in the first row represent the evidence in favor of type $2_2[4]$ being easier to learn than type $2_2[0]$. The plots in the second row represent the t test result that was closest to providing evidence for the null hypothesis of no statistical difference between type $2_2[2]$-II and the other three non-trivial structure types ($2_2[1]$, $2_2[2]$-I, and $2_2[3]$). The plots in the third row represent the lone difference between structure types $2_2[1]$, $2_2[2]$-I, and $2_2[3]$. The plots were created using JASP (JASP Team 2016)

References

Archer EJ (1962) Concept identification as a function of obviousness of relevant and irrelevant information. J Exp Psychol 63(6):616–620

Bourne LE (1970) Knowing and using concepts. Psychol Rev 77(6):546–556

Bourne LE, Guy DE (1968) Learning conceptual rules. II: the role of positive and negative instances. J Exp Psychol 77(3):488–494

Bourne LE, Ekstrand BR, Montogomery B (1969) Concept learning as a function of the conceptual rule and the availability of positive and negative instances. J Exp Psychol 82(3):538–544

Bruner JS, Goodnow JJ, Austin GA (1956) A study of thinking. Transaction Publishers, New York

Bulgarella RG, Archer EJ (1962) Concept identification of auditory stimuli as a function of amount of relevant and irrelevant information. J Exp Psychol 63(3):254–257

Byrne RM, Johnson-Laird PN (2009) ‘If’ and the problems of conditional reasoning. Trends Cogn Sci 13(7):282–287

Conant MB, Trabasso T (1964) Conjunctive and disjunctive concept formation under equal-information conditions. J Exp Psychol 67(3):250–255

Dijkstra S, Dekker PH (1982) Inference processes in learning well-defined concepts. Acta Physiol (Oxf) 51:181–205

Dobson DJG, Dobson KS (1981) Problem-solving strategies in depressed and nondepressed college students. Cognit Ther Res 5(3):237–249
Feldman J (2000) Minimization of Boolean complexity in human concept learning. Nature 407:630–633
Feldman J (2003) A catalog of Boolean concepts. J Math Psychol 47:75–89
Feldman J (2006) An algebra of human concept learning. J Math Psychol 50:339–368
Goode RL (2001) Auditory physiology of the external ear. In: Jahn AF, Santos-Sacchi J (eds) Physiology of the ear, 2nd edn. Singular, San Diego, pp 147–160
Goodwin GP, Johnson-Laird PN (2011) Mental models of Boolean concepts. Cogn Psychol 63:34–59
Halberstadt J, Sherman SJ, Sherman JW (2011) Why Barack Obama is black: a cognitive account of hypodescent. Psychol Sci 22(1):29–33
Haygood DH (1965) Audio-visual concept formation. J Educ Psychol 56(3):126–132
Haygood RC, Bourne LE (1965) Attribute- and rule-learning aspects of conceptual behavior. Psychol Rev 72(3):175–195
Higonnet RA, Grea RA (1958) Logical design of electrical circuits. McGraw-Hill, New York
Hull CL (1920) Quantitative aspects of the evolution of concepts. Psychol Monogr 28(1):1–86
Khemlani S, Orenes I, Johnson-Laird PN (2014) The negations of conjunctions, conditionals, and disjunctions. Acta Physiol (Oxf) 151:1–7
Kruschke JK (1992) ALCOVE: an exemplar-based connectionist model of category learning. Psychol Rev 99(1):22–44
Kurtz KJ, Levering KR, Stanton RD, Romero J, Morris SN (2012) Human learning of elemental category structures: revising the classic result of Shepard, Hovland, and Jenkins (1961). J Exp Psychol Learn Mem Cogn 39(2):552–572
Lordahl DS (1961) Concept identification using simultaneous audio and visual signals. J Exp Psychol 62(3):283–290
Luce RD (1959) Individual choice behavior: a theoretical analysis. Wiley, New York
Miskiewicz A, Rakowski A (2012) A psychophysical pitch function determined by absolute magnitude estimation and its relation to the musical pitch scale. J Acoust Soc Am 131:987–992
Murphy GL (2002) The big book of concepts. MIT Press, Cambridge
Neisser U, Weene P (1962) Hierarchies in concept attainment. J Exp Psychol 64(6):640–645
Newstead SE, Griggs RA, Chrostowski JJ (1984) Reasoning with realistic disjunctives. Q J Exp Psychol 36A:611–627
Nosofsky RM (1984) Choice, similarity, and the context theory of classification. J Exp Psychol Learn Mem Cogn 10(1):104–114
Nosofsky RM (1986) Attention, similarity, and the identification-categorization relationship. J Exp Psychol 115(1):39–57
Nosofsky RM, Johansen MK (2000) Exemplar-based accounts of "multiple-system" phenomena in perceptual classification. Psychon Bull Rev 7(3):375–402
Nosofsky RM, Palmeri TJ (1996) Learning to classify integral-dimension stimuli. Psychon Bull Rev 3(2):222–226
Nosofsky RM, Gluck MA, Palmeri TJ, McKinley SC, Glauthier P (1994) Comparing models of rule-based classification learning: a replication and extension of Shepard, Hovland, and Jenkins (1961). Mem Cogn 22(3):352–369
Paivio A (1971) Imagery and verbal processes. Holt, Rinehart, and Winston, New York
Paivio A (1986) Mental representations: a dual-coding approach. Oxford University Press, Oxford
Pishkin V, Shurley JT (1965) Auditory dimensions and irrelevant information in concept identification of males and females. Percept Mot Skills 20:673–683
Rehder B, Hoffman AB (2005) Eyetracking and selective attention in category learning. Cogn Psychol 51:1–41
Reznick JS, Ketchum RD, Bourne LE (1978) Rule-specific dimensional interaction effects in concept learning. Bull Psychon Soc 12(4):314–316
Shepard RN (1987) Towards a universal law of generalization for psychological science. Science 237:1317–1323
Shepard RN, Hovland CI, Jenkins HM (1961) Learning and memorization of classifications. Psychol Monogr General Appl 75(13):1–42
Smith ER, Zarate MA (1992) Exemplar-based model of social judgment. Psychol Rev 99(1):3–21
JASP Team (2016) JASP (Version 0.8.0.0) [Computer software]
Vigo R (2006) A note on the complexity of Boolean concepts. J Math Psychol 50:501–510
Vigo R (2009) Modal similarity. J Exp Theor Artif Intell 21:181–196
Vigo R (2013) The GIST of concepts. Cognition 129(1):138–162
Vigo R (2014) Mathematical principles of human conceptual behavior. Routledge, New York
Vigo R, Allen C (2009) How to reason without words: inference as categorization. Cogn Process 10:77–88
Vigo R, Basawaraj B (2013) Will the most informative object stand? Determining the impact of structural context on informativeness judgments. J Cogn Psychol 25(3):248–266
Vigo R, Barcus M, Zhang Y, Doan C (2012) On the learnability of auditory concepts. Paper presented at the 146th meeting of the Acoustical Society of America, Kansas City, MO
Walls RT, Rude SH, Gulkus SP (1975) Model and observer learning of low, medium, and high level concepts. Psychol Rep 37:671–675