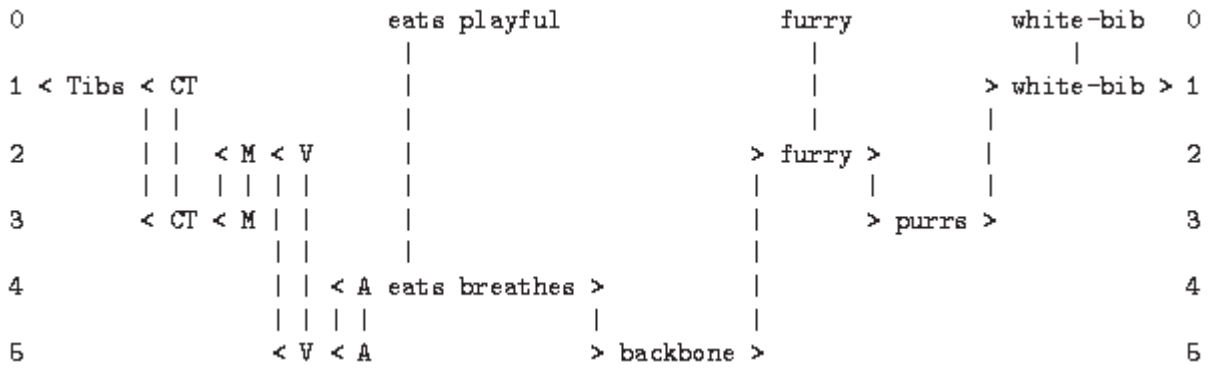


Figure 2 (Wolff). A multiple alignment showing how an unknown entity may be recognized or categorized at multiple levels of abstraction.

rules or similarity, but we can recognize these concepts in the way the system operates in processing natural language, in categorization of an unknown entity, in various styles of reasoning, and in the unsupervised learning of new structures.

NOTE

1. Notice how "b o y" and "g i r l" have been assigned to the same category ("%2"), in the spirit of distributional linguistics.

Author's Response

Preferring Rules to Similarity: Coherence, goals, and commitment

Emmanuel M. Pothos

Department of Psychology, University of Crete, Rethymnon 74100, Greece.
pothos@psy.soc.uoc.gr **http://www.soc.uoc.gr/pothos/**

Abstract: This response to the open peer commentary discusses what should be the appropriate explanatory scope of a rules versus similarity proposal and accordingly evaluates the Rules versus Similarity one. Additionally, coherence, goals, and commitment are presented as inferential notions, fully consistent with the Rules versus Similarity distinction, that allow us to predict when Rules would be preferred to Similarity.

R1. What is the aim of a rules versus similarity proposal?

Evaluating a rules versus similarity (r-s) proposal is a complex issue, not least because, as the commentaries show, there is no universal agreement of what the explanatory scope of such a proposal should be. In his commentary, **Duch** examines whether brain neurons might be seen as computing rules. But clearly an r-s proposal aims to characterize psychological behavior (neuronal collective activity), not biological or neurological functions. Now, as Marr (1982) discussed, brain function can be investigated at different explanatory levels and explanations at different levels must be consistent with each other. However, our current knowledge of neuron properties offers weak constraints in preferring one notion of rules as opposed to another (**Campion**).

The second point to address is how *much* of psychological behavior we should aim to explain within a model of r-s. Unfortunately, psychological behavior is not a one-dimensional system that can be neatly divided into non-overlapping units for analysis. **Campion** goes through an example of real-life inference in order to make the point that the "subtle psychological mechanisms" supporting the inference are obscured in a characterization in terms of rules or similarity. However, an r-s proposal is most certainly not meant to constitute a universal model of cognition. It goes without saying that beyond rules and similarity, several other characterizations for a cognitive process would be possible.

How detailed should an r-s proposal be? Several researchers, including myself in the target article and some of the commentators (**Duch**; **Hampton**; **Markman**, **Blok**, **Kim**, **Larkey**, **Narvaez**, **Stilwell**, **& Taylor** [henceforth **Markman et al.**]), have pointed out that, unconstrained, the general notions of rules and similarity are so flexible so as to enable arbitrary characterizations of performance. Therefore, the least we should require of an r-s proposal is that it enables a principled distinction of cognitive operations into rules and similarity. **Dominey** goes further to argue that such a proposal must also be successfully implemented in a computational model. However, the lack of a suitable computational model for an r-s proposal should not determine exclusively the validity of the proposal, insofar that our limitations in proposing adequate computational models for cognitive theories cannot always inform us of the validity of the theories (but of course it has to be shown that the cognitive processes postulated in a theory are computable by the brain; cf. Newell 1990).

An r-s proposal should aim to clarify which operations we would like to consider rules and which similarity. But what would be the basis of such an endeavor? Clearly, we should start with situations that most psychologists would informally want to characterize as rules (e.g., deciding that a number is even) and empirical data that are generally accepted as showing rule performance (for example, recognizing that a sentence such as "colorless green ideas sleep furiously" is grammatically correct). Some of the characteristics of these situations and empirical data would serve as the *assumptions* of an r-s proposal, on the basis of which we would try to predict the remaining characteristics (and, of course, r-s dissociations more generally). In doing this, we can recognize two facts of scientific reality: first, we should aim to formulate our theory with the least number of as-

https://doi.org/10.1017/S0140525X05470016 Published online by Cambridge University Press

sumptions, and second, which characteristics serve as assumptions is a somewhat arbitrary choice. Therefore, the predictions of one r-s model could be the assumptions in another, and characteristics that are central in some proposals may be obscured in others (cf. **Brooks & Hannah**). For example, **McMurray & Gow** imply that we should consider an operation as similarity depending on the gradedness of the operation. Within the Rules versus Similarity (R-S) proposal, the definitional centrality of gradedness is rejected, even if in practice most Similarity operations would be graded and most Rule ones discrete (as in the target article, capitalized "Rules" and "Similarity" refer to the present proposal). To conclude with the above points, an r-s proposal is considered partly a definitional issue (as discussed further in sect. R4.2). In this sense, **Hahn**'s criticism that "Structural relationships between old and new items have always been viewed as *evidence* for rules or similarity, but the present proposal . . . turns what was evidence of the conceptual contrast between rules and similarity into their definition" (para. 8), simply illustrates the definitional choices made in the R-S proposal. Rather than denounce these choices a priori, we should aim to examine how well they serve the explanatory purpose of an r-s proposal.

The aim of the target article was to provide a formulation of the r-s distinction in a way that would enable consistent and unambiguous characterizations of cognitive processes as rules or similarity. However, **Sloman** correctly observes that "More challenging would be a theory that . . . explains why the phenomena of thought cluster [in terms of rules of similarity] as they do" (cf. **Heit & Hayes**). The primary aim of the present response is to discuss the implications of the R-S proposal with respect to when we would expect Rules behavior as opposed to Similarity.

R2. Assumptions: Clarifying the Rules versus Similarity formulation

R2.1. Relevant properties and representation

Some commentators (**Hahn**, **Heit & Hayes**, **Lupyan & Vallabha**) have questioned the appropriateness of the representational assumptions made in the R-S proposal. For example, Hahn claims that the notion of uniquely common properties is not robust and Lupyan & Vallabha observe that object representations can develop dynamically as a result of experience (Goldstone 1998).

Starting with **Lupyan & Vallabha**, the R-S proposal is consistent with their observation. The proposal requires that when we process an object it is possible to specify its properties relevant to the process. The specification of this representation depends on the particulars of the process. To use **Smith**'s example: If two obscured pictures were compared, the psychological representation of the pictures would correspond only to the shown parts of the pictures (so the comparison would be Similarity, if it takes into account all the shown information). Analogously for **McMurray & Gow**, an R-S characterization does not depend on whether a process involves one or more properties, but, rather, on the number of properties involved out of all the relevant properties. For example, are there alternative voicing perception operations that depend on multiple stimulus dimensions? If there is always only one dimension then the operation is Similarity. Note that the representation in itself does not necessarily (sect. 3.1) determine whether a Rules process will be preferred to a Similarity comparison; if it did, the R-S proposal would be circular (**Sloman**).

I agree with **Hahn**'s observation: The lack of an adequate theory for representation is one of the important current shortcomings in psychology, and the R-S proposal does not claim to address it. However, the corresponding assumptions made in the R-S proposal are analogous to the assumptions made in models of categorization generally (e.g., Nosofsky 1991). Despite the acknowledged inadequacy of these assumptions, such models work impressively. Therefore, the assumptions cannot be entirely inappropriate, and are plausibly approximate to whatever more compelling approaches would be forthcoming in the future. With respect to the R-S proposal, the enumeration of relevant properties is done by considering the properties that would be (uniquely) required under each *reasonable* possible version of the operation of interest on a target representation. In this way, when, for example, inflecting a verb in English, we could be taking into account its phonology, its status as regular or irregular, and its semantics.

Davidoff discusses results that show a dissociation between the perception of abstract and concrete properties. The R-S proposal is concerned not with how the properties involved in a representation arise, but, rather, with whether they are differentially considered in object judgments. For example, would it be the case that a Rule operation would be behaviorally different if it is based on a single abstract property instead of a single concrete one? According to the R-S proposal the answer is no, and Davidoff's results suggest no reason to revise this answer. Note that Rules will be more abstract than Similarity in the sense that Rules are operations on a target representation that involve few properties of the representation (**Bailey**).

R2.2. Computations of Rules *versus* Similarity

The discussion below has two parts. We first clarify the assumptions about Rules processes versus Similarity ones that are integral in the definition of the proposal. We then consider some implications.

Consider a task whereby participants are presented with three letter strings: MSSSSSV, MSSV, and MSSSSV. Participants subsequently decide that new string MSSSV is compatible with the training ones. **Hahn** observes that such a judgment could equally be considered Similarity (the features of the new string overlap with the features of the old ones) or Rules (old items considered to conform to the category "strings start with an M, end with a V, and have S's in between"). Strictly speaking, an R-S characterization is forthcoming only for an assembly of judgments. When a diverse enough set of judgments is examined (e.g., how would participants characterize item MSSSSSSSV?), then we could conclude: "Well, it turns out the person was using Similarity," or "It turns out the person was using Rules." Accordingly, our characterization of what participants do may change as the test set is expanded. However, contrary to Hahn's view, such an implication is hardly inappropriate: It means that the more information we have about people's behavior, the more accurate our characterization of the corresponding psychological process. Indeed, dissatisfaction with the r-s distinction is possibly due to researchers' inclination to define the distinction in terms of inadequate empirical measures (**Dulany**). Finally, note that the letter

https://doi.org/10.1017/S0140525X05470016 Published online by Cambridge University Press

string discussion is equivalent to the one in the target article for conditionals. However, as **Arló-Costa**'s commentary shows, we need to be careful about how we interpret performance with conditionals. In this respect, the discussion in the target article should be simply seen as an illustration of how an R-S analysis could proceed in reasoning.

Commentator **Reber** asks us to consider participants who are told exactly how to compute the associative chunk strength of the new string (a Similarity measure) in order to decide whether the new string is consistent with the old ones or not. Reber claims that for participants to carry out such a computation they must be using rules; however, by the R-S proposal, the outcome of the computation would still be (inappropriately) considered Similarity. I agree that the individual computations required to compute associative chunk strength would correspond to Rules (cf. **Hampton**, **Smith**). However, an overall characterization of the result of the computation in terms of Rules or Similarity is no longer appropriate, since the string classification is no longer a single process (or a set of integrated processes – see the discussion for **Lupyan & Vallabha** later on), but the result of a set of independent processes. Reber further suggests that the use of a threshold should be considered a rule. However, evaluating the outcome of a computation on a representation in terms of a threshold can equally correspond to Similarity or Rules, depending on whether most or few of the properties of the representation are taken into account (cf. **Duch**). Indeed, in psychology, thresholds are not generally associated with psychological rules: For example, the operation of exemplar categorization models can be seen as involving thresholds, but no one would argue that such models reflect rules.

Most cognitive processes are the result of an assembly of integrated tasks and influences. For example, the past-tense inflection of a verb might depend on processing of its phonology, semantics, and so forth. **Lupyan & Vallabha** observe that it is not appropriate to characterize each of these influences as Rules or Similarity individually and thus question the practical utility of the proposal. In such cases, the meaningful way to apply an R-S characterization is with respect to the aggregate result of the processing on a target representation, by considering together all the components of the processing. For example, is it the case that the component parts of a past-tense inflection collectively result in a process that involves most of the relevant properties of the verb, or only a small subset? By contrast, when a set of tasks is carried out independently of one another, each task should be individually examined as to whether it is a Rule or Similarity (**Reber**). Note that these considerations appear relevant to any r-s proposal.

Sloman notes that in a rules system we typically require consistency, in an associative system we do not, and that this is a qualitative difference between rules and similarity that cannot be accounted for within the R-S framework. I think that an obvious inconsistency in Similarity will bother us as much as an obvious inconsistency in a Rule. For example, I will never believe that the object in front of me is both red and not red (judgment depends on one feature; a possible Rule inconsistency). Likewise, I will never believe that the object in front of me is both a chair and not a chair (judgment depends on several features; a possible Similarity inconsistency). So let's just assume that the cognitive system would avoid obvious inconsistencies. Now, for a Rules system, the consistency requirement is more stringent than for a Similarity one, because Rules are more certain and there is less ambiguity as to whether they apply or not. By contrast, Similarity judgments depend on most properties of the representations that are compared, so that two objects will be Similar (or Dissimilar) in several different ways. In other words, there would be more ambiguity in Similarity judgments and therefore the need for consistency in such judgments would be less pronounced.

R2.3. Memory involvement

In this section we continue our consideration of Rules computations versus Similarity ones, but with an emphasis on the way memory supports these operations.

Marcus points out that rules are computational operations on a target instance that may have arbitrary algebraic complexity; by contrast, similarity is a metric that "depends on . . . the extent to which two inputs share representational material." An analogous point is made by **Bailey**, who suggests that rules are operations on a single target representation, which is checked with respect to whether the rule applies or not. By contrast, similarity is a comparison between two representations. **Smith** also assumes that in similarity comparisons an object is compared to a representation holistically, whereas in rules operations the features of the target representation must be examined explicitly to determine whether the rule applies.

Let's start with Similarity operations, because there appears to be less controversy about how these are carried out. Suppose we have a target representation, say a particular telephone ("the telephone John has"), and we wish to determine how Similar it is to other telephones in our memory. This operation must (effectively) involve some comparison of the properties of telephones in our memory and the properties of the target representation. Consider now a Rules operation, whereby we try to determine whether the target representation is a telephone or not: Our Rule for telephones is that "a telephone is any device which conveys sound over a distance." There must be some representation for the Rule that is activated in order to apply the Rule. Then, we need to examine whether the (few) properties involved in the Rule representation match the (correspondingly few) properties of the target representation. This is just like the Similarity process above, but in a reduced scale: Instead of comparing all the features between (say) the two representations, I only compare a small subset of these. Additionally, the comparison of features between the Rule representation and the target representation must be based on a process analogous to the comparison of features for a Similarity operation, because in both cases (effectively) the basic elements of the cognitive operations are the same: pairwise comparisons of features (cf. **Brooks & Hannah**, **Dulany**, **Hampton**).

Now, with respect to memory, it is typically assumed that Rules correspond to knowledge that is somehow independent of particular instances. Within the R-S proposal this does not have to be the case (contra **Bailey**). A rule operation for a set of instances specifies a restricted representation for the instances, so that they can differ only in terms of the properties involved in the Rule. For example, when I consider the Rule for regular past-tense inflection in English, from the perspective of this operation all regular verbs are nearly identical. Accordingly, my knowledge of regular inflection is supported by many exemplars, in the

https://doi.org/10.1017/S0140525X05470016 Published online by Cambridge University Press

same way that my knowledge of different kinds of irregular inflection is supported by exemplars. In fact, it looks like my knowledge of a Rule fact would often be supported by *more* exemplars than my knowledge of a Similarity fact (cf. Hintzman 1986). For example, by the present account, regular inflection in English involves more instances than irregular inflection. Therefore, regular inflection would be predicted to depend less on the intactness of memory representations compared to irregular inflection (cf. Joanisse & Seidenberg 1999; Miozzo 2003). Moreover, even if Rules operations are less graded than Similarity ones (because of certainty), Rules operations would still display some gradedness. Indeed, Albright and Hayes (2003) have shown gradedness in acceptability of past tense forms for nonce English verbs for both regular and irregular verbs.

To conclude, there is little doubt that rules and similarity computations are different and also that knowledge of rules is supported in memory in a different way from knowledge of similarity, as **Bailey**, **Marcus**, and **Smith** have pointed out. The R-S proposal provides an alternative perspective for this issue by showing how some of these differences can result from minimal assumptions about how Rules differ from Similarity in degree and not qualitatively.

R3. Implications – predicting Rules behavior as opposed to Similarity

R3.1. Coherence

In this section we are interested in when Rules would be spontaneously preferred to Similarity. **Heit & Hayes** point out that in reasoning there is a lot of evidence for a dual process system, whereby an analogy (Similarity) mode appears to be the default and a rules (Rules) mode is possible after additional effort. Recent empirical investigations corroborate this view. For example, Schroyens et al. (2003) found that under time constraint participants preferred logically invalid conclusions to valid ones, so that it looks as if reasoning on the basis of (logical) rules is not the default. Note that Rules performance in reasoning implies that few of the properties of a problem (its abstract logical structure) are taken into account in reaching a conclusion. It is not the case that an (overall) similarity heuristic for all reasoning processes is advocated (**Arló-Costa**).

Heit & Hayes (and **Ashby & Casale**) also observe that in spontaneous categorization experiments participants typically produce single-dimensional sorts. That is, participants ignore most of the dimensions of the objects and derive a grouping for the objects on the basis of a single dimension. Therefore, it looks like Rules is the default in spontaneous categorization, even if that's not always the case (Pothos & Chater 2002).

To paraphrase **Heit & Hayes**, why does the number of features vary (cf. **Sloman**)? Why is it the case that in reasoning the default appears to be Similarity but in spontaneous categorization experiments the default appears to be Rules? According to the R-S proposal, this is a problem of category coherence (Murphy & Medin 1985). We spontaneously recognize certain groupings of objects as more intuitive than others. For example, the grouping of all instances of a "cat" into the same concept is highly intuitive, however, this is hardly the case for the grouping of "the Eiffel tower, dolphins, and the moon." Now, coherence between the members of different categories could be established in different ways. For example, consider the category of chairs: The instances of this category appear to cohere together because they are all highly Similar to each other (any two instances of the category will share many characteristics). By contrast, the instances of the category of pens can vary greatly in terms of most of their properties. What they have in common is their function, so that the instances of the category of pens appear to cohere together in terms of a Rule. Correspondingly, with a model of category coherence (e.g., Pothos & Chater 2002) we can predict whether a set of objects would be spontaneously perceived in terms of a Rule or Similarity, depending on which possibility enhances the coherence amongst the objects more.

In this way, it appears that in reasoning problems the spontaneous predisposition would be to take into account all the properties of the problems – both their pragmatic content and their logical structure. Therefore, in looking for conclusions that depend only on their logical structure (see sect. R3.2, Goals) we have to use selective attention to focus on some of the problem properties (logical structure) and suppress others (pragmatic content). The use of selective attention implies that Rules in logical reasoning would be more effortful (**Heit & Hayes**) and explicit (**Smith** – this is the reason that Rules operations are typically explicit; sect. R4.2). Analogous considerations apply to spontaneous categorization experiments. Note that this conclusion can be seen as partly confounding the correlations between IQ and ability in logical reasoning, as IQ measures have also been shown to correlate with ability in attentional tasks (e.g., Nijenhuis & van der Flier 2002). Furthermore, this discussion should clarify that the R-S proposal is more consistent with dual process models of reasoning (cf. Heit & Hayes); the fact that Rules and Similarity are assumed not to differ qualitatively far from implies that they do not differ in many and important respects. With respect to single process models of reasoning (e.g., Oaksford & Chater 1994), I suggest that such models often have an explanatory objective that is different from that of characterizing the reasoning process in terms of rules or similarity.

R3.2. Goals

Category coherence would determine the spontaneous bias for processing a set of objects. Clearly, our goal associated with the objects may not be consistent with this spontaneous bias, in which case we would have to employ selective attention to process the objects in a particular way. Goals can be explicit, as in the reasoning example above. We can also identify what we could call, following **McMurray & Gow**, system goals: We can (minimally) assume that the cognitive system will aim to be as efficient as possible while avoiding errors. Below we present some corresponding implications for the R-S proposal.

Let's first consider whether second language (L2) acquisition involves Rules or Similarity (**Robinson**). If there is a lot of superficial mismatch between L2 representations and L1 (mother tongue) representations, in trying to make sense of L2 stimuli it would be reasonable to focus on whichever few aspects of L2 stimuli are analogous to L1 ones. In other words, the cognitive system would be trying to encode L2 stimuli in terms of Rules that are guided by L1 knowledge. Likewise, if the structures of L1 and L2 share many commonalities, encoding an L2 stimulus could take place by matching many of its properties to a corre-

https://doi.org/10.1017/S0140525X05470016 Published online by Cambridge University Press

sponding L1 stimulus (Similarity). Accordingly, following Robinson, in second language acquisition we can emphasize learning on the basis of Rules or Similarity, depending on the characteristics of L2 and the relation of L2 to L1.

Wolff's model can illustrate the above. The model searches for the briefest code for new instances in terms of previous ones (Chater 1999). For example, new stimulus MSSSV could be encoded with stimulus MSSS from my memory or pattern MS, abstracted across many stimuli in my memory. MSSS provides more complete coverage of MSSSV, but it would (generally) be associated with a low frequency (hence high codelength). MS covers MSSSV very partially, but it would have a high frequency and hence low codelength. Information theory can help decide which will be preferred and so, whether a Rule (choosing MS to encode the new instance) would be preferred against Similarity (choosing pattern MSSS).

McMurray & Gow point out that "Research on speech perception demonstrates that the degree to which speech categories are gradient (or similarity-based) is a function of the utility of within-category variation for further processing" (Abstract). I endorse the general point McMurray & Gow make. The cognitive system ignores information all the time. For example, we rarely take into account orientation information in processing a pen, presumably because historically such information is useless. Likewise, whether the properties in a representation are given equal salience (Similarity) or not, is plausibly influenced by (statistical) expectations of how the representation will be utilized in the future (since in this way future needs can be anticipated and dealt with more expediently). The above discussion leads to **Diesendruck**'s notion of commitment.

R3.3. Commitment

Diesendruck's main point is that a fundamental distinction between rules and similarity is that rules require commitment. In other words, "rules have a motivational and stabilizing force that derives from people's commitment to rules. . . . Commitment is what licenses leaps of faith, allowing categorizers to draw inferences even about novel categories and properties in familiar domains" (Diesendruck, para. 2). Diesendruck suggests that commitment depends collectively on our experiences and knowledge base. The commitment point is echoed in **Dulany**'s commentary as well. According to Dulany, we form explicit beliefs about which rules (or similarity comparisons) will allow us to decide how novel instances should be categorized.

Suppose we initially represent an object in terms of features ABCDE and we want to process it in terms of a Rule that involves feature A. However, we may later revise the object's representation and identify new feature A?, which we would want to consider more important than A (**Lupyan & Vallabha**). Plausibly, the cognitive system will not be "convinced" to attend to a single feature in a representation and suppress the rest, unless it is "convinced" that additional processing would not lead us to identify new, potentially important, features (equivalently, that in additional processing of the object certain few features would always be more significant than the rest). This "conviction" will likely depend on general knowledge and appears equivalent to **Diesendruck**'s commitment. Moreover, the above can be rephrased in the context of **McMurray & Gow**'s discussion in the previous section. For a domain of stimuli for which even though a single aspect of their representation is important, other aspects of the stimuli may also be useful in the context of the same or a similar operation in the future, suppression of the irrelevant properties would be less forthcoming.

The above interpretation of commitment is consistent with the R-S proposal in that it shows circumstances that would enhance Rules performance as opposed to Similarity but does not show rules that have to be qualitatively different from similarity.

R4. Evaluating the Rules versus Similarity proposal

R4.1. Coverage of existing data – do we need separate rules versus similarity systems?

The most direct prediction of the R-S proposal is that Rules operations are not qualitatively different from Similarity ones. **Ashby & Casale** present a series of studies to show separate (qualitatively distinct) systems of rules and similarity. Ashby & Casale argue that these results are inconsistent with any unified r-s proposal, including the R-S one. I agree that these results show separate systems, but these do not (necessarily) correspond to rules and similarity, rather (plausibly) to implicit and explicit learning modes. Implicit versus explicit modes of learning may be confounded with r-s learning in experiments such as the ones discussed by Ashby & Casale in the following way: We can generally assume that the default mode of processing for objects such as those in the experiments they discuss would be Similarity, in that most of the object properties would be initially perceived as having equivalent salience. Therefore, participants required to process the objects in terms of a Rule, would have to selectively attend to some of the object features (cf. **Brooks & Hannah**, **Smith**). The intentionality required for a selective attention process implies that this process would be explicit (sect. R4.2). By contrast, when no selective attention is needed, the stimuli are plausibly processed implicitly (where implicitly is used here in the sense of incidental; cf. artificial grammar learning). So, the behavioral differences discussed by Ashby & Casale could be partly due to differences in rules and similarity and partly due to differences in implicit and explicit learning. Given that it is possible that implicit and explicit learning could correspond to separate systems, I suggest that without further insight about which performance differences are due to a contrast between rules and similarity and which due to a contrast between implicit and explicit learning, it is not possible to claim that rules and similarity necessarily correspond to separate systems. The R-S proposal is not about there being a single R-S continuum to characterize all, but, rather, that we can interpret rules and overall similarity in the *same* system as extremes in an R-S continuum (cf. Brooks & Hannah, **Sloman**).

An analogous point applies to the imaging results **Dominey** and **Smith** discuss. Smith notes that in the condition requiring an explicit rule there was activation in the areas mediating selective attention and inhibition, but not in the similarity condition. This finding is consistent with the R-S proposal, as Rules operations require selective attention when Similarity is the default. More generally, in evaluating the implications for the r-s debate from such results, the perspective of Juola and Plunkett (1998) is useful.

https://doi.org/10.1017/S0140525X05470016 Published online by Cambridge University Press

Using neural networks, they showed that dissociation results can correspond to extreme aspects of the operation of a single system, rather than indicating separate systems.

R4.2. Do rules have to be explicit?

Many commentators believe that rules have to be explicit (**Cleeremans & Destrebecqz**, **Hampton**, and **Smith**). In Smith's model the first defining characteristic of rules performance is that rules are explicitly represented and the second defining characteristic is that rules can be verbalized. Cleeremans & Destrebecqz emphasize this point even more by titling their commentary "Real rules are conscious."

If we restrict rules to consciously accessible entities then we have difficulty in characterizing the operation via which we can instantly recognize the sentence "colorless green ideas sleep furiously" as grammatically correct. Because this sentence is meaningless, there are no instances in our experience we could use to classify it as grammatical on the basis of similarity. Additionally, we do not explicitly apply grammar rules – many of us hardly know what these are (in other words, no metaknowledge of grammar rules; cf. **Cleeremans & Destrebecqz**). We have to postulate that we use rules knowledge that is not explicit. **Hampton** suggests that in such cases "the person can be said to be following a rule, but this is not evidence that the rule itself is represented in the part of the mind/brain directing the behavior" (cf. Cleeremans & Destrebecqz, **Smith**).

Objections to the psychological status of nonexplicit rules appear to boil down to the following issue: Unless we know we are using a rule, there is no evidence that a rule is being used. A fundamental aspect of the R-S proposal is that, regardless of whether rules are implicit or explicit, it is not possible to prove their existence, because their specification is partly a definitional issue. This perspective can be justified by generalizing **Smith**'s profound insight: An explanatory notion in psychology, such as a rule, is like a concept that groups together certain psychological behaviors. Subsequently, by studying the behaviors in, for example, the rules category, we can try to determine whether it is possible to characterize them in some parsimonious way. However, as is the case with the psychological process of categorization, there is often considerable arbitrariness about which instances/behaviors will be associated with a theoretical concept, which then affects the characterization of the concept. In this way, I suggest that rather than seeking to prove an r-s proposal, we should examine whether it allows for a more convenient (from an explanatory perspective) categorization of psychological processes (**Heit & Hayes**).

Cleeremans & Destrebecqz, **Hampton**, and **Smith** argue that we ought to categorize differently behaviors based on explicit rules from behaviors that could be described in terms of (implicit) rules. However, in both cases, rules reflect collective neuronal activity, so they do not exist in any way other than as emergent properties of neurons. So, from such a perspective, the brain is a physical system, just like a river, whose properties we seek to characterize in some mathematical way (Hampton). Moreover, neural networks do not have metaknowledge of rules, but arguably they have no metaknowledge of any aspect of their knowledge (Cleeremans & Destrebecqz; cf. Dienes & Perner 1999). Additionally, while we may explicitly believe that we are using a rule, our metaknowledge may be misleading and the actual psychological process could have an alternative basis (cf. A.R. Reber et al. 1985).

Having said the above, it has to be appreciated that most of our intuitions of rule following correspond to explicit, verbalizable rules. Regardless of whether a particular r-s proposal includes in its definition rules explicitness (and associated features), these intuitions must be explained. The R-S proposal predicts that most Rules would be explicit and require selective attention if the default mode of processing for the stimuli we encounter is Similarity. As noted in section R3.1 above, whether this is the case is an issue of category coherence. Intuitively, however, it would seem reasonable to assume that when the cognitive system encounters a novel object it would not spontaneously ignore most of its properties, so that Similarity is indeed the default.

To summarize, if we define Rules as operations that involve few of the properties of a target representation, then selective attention and explicitness for rules operations are implications of this definition in (what appear to be) the most common rules situations.

R4.3. Do we need rules (or similarity)?

Several commentators question the validity of the r-s debate (**Calvo Garzón**, **Hampton**, **Markman et al.**, **Smith**, **Vilarroya**, **Wolff**). For example, Markman et al. discuss three performance dimensions along which rules and similarity have been traditionally distinguished, to argue that distinctions along these dimensions are misguided. They suggest that researchers should emphasize the study of the performance dimensions themselves and worry less about r-s characterizations. Vilarroya articulates these concerns in a more forceful way, pointing out that "the progress of similarity-based theories of learning and categorization will reduce the area of influence of rules to a small corner." According to this logic, maybe we should just abandon rules as an explanatory concept in psychology.

It must be appreciated that unconstrained rules and similarity are indeed vacuous in that we can always conceive of alternative notions of rules (or similarity) to describe arbitrary patterns of performance. Indeed, this has been the motivation for the R-S proposal – and another way of saying that the specification of rules and similarity is partly a definitional issue (sect. R4.2). If the notions of rules and similarity are restricted in a suitable way, then the problems raised by **Hampton**, **Calvo Garzón**, and **Markman et al.** disappear. For example, with respect to Markman et al., Rules cannot be distinguished from Similarity in terms of whether they involve abstract properties or not. But this insensitivity to abstraction of properties is not arbitrary. Rather, it is an explicit choice in the definitional specification of the model. In alternative r-s models abstraction may have a more central role (e.g., cf. **Brooks & Hannah**). As noted, which proposal is preferred should be an issue of explanatory convenience.

Calvo Garzón further suggests that the information-processing framework for cognition, within which the rules versus similarity debate has been formulated, may be inappropriate. His discussion is a useful reminder that sometimes we are so immersed in the specifics of a particular debate that we lose sight of plausible alternative frameworks for describing cognition. It is beyond the scope of this pa-

https://doi.org/10.1017/S0140525X05470016 Published online by Cambridge University Press

per to examine how the information processing framework might measure up to alternatives. Nevertheless, there has to be a sense in which the initial basis of cognitive models is our introspective understanding of cognition (sect. R1). As discussed in the target article, we do have a strong intuitive sense in which sometimes we are applying a rule while in other cases we are carrying out similarity comparisons (and indeed, some models of r-s performance are derived almost entirely from this intuition; e.g. **Smith**). Accordingly, there has to be a specification of theories of cognition that enables some role for rules and similarity (cf. **Vilarroya**). Moreover, from a purely theoretical perspective, it would be useful to retain a distinction between rules and similarity to the extent that the processes considered rules differ interestingly from the processes considered similarity. The data discussed in the target article suggest this to be the case.

In conclusion, a priori there is every reason to expect that an r-s distinction would be important in psychological theory. It is argued that dissatisfaction with existing r-s proposals is either due to underspecification (so that researchers worry that the distinction is vacuous; cf. **Dulany**), or because they do not afford a useful/interesting differentiation of cognitive processes. The above observations suggest that rather than dismissing the r-s distinction altogether, we should seek to develop alternative r-s approaches that address the shortcomings of previous ones.

R5. Summary

Operations on a target representation can be characterized in terms of Rules or Similarity depending on whether few or most of the properties of the target representation are involved. In this way, a distinction between Rules and Similarity is understood as a distinction in degree, not quality, so that Rules and Similarity are extreme aspects of the same (similarity) operation. In other words, rules are certainly not dismissed in favor of similarity (cf. **Markman et al.**); rather, it is suggested that the relation between rules and similarity is such that we need not postulate a separate rules system from a similarity one.

The R-S continuum is only one possibility amongst many in how to characterize processes as rules and similarity. Following **Smith**, it is suggested that such explanatory notions are like categories that group together certain behaviors. Accordingly, there is some definitional arbitrariness regarding how rules and similarity are specified – in other words, regarding which behaviors we group in the category of "rules" and which in the category of "similarity." Despite this arbitrariness, there are certain characteristics that we would generally wish to associate with rules and others we would want to associate with similarity.

The claim in the R-S proposal is that these characteristics can be predicted if we define rules as Rules and similarity as Similarity. Therefore, Rules are predicted to be more certain and less graded than Similarity, and in many cases Rules would be explicit, verbalizable operations that require selective attention (**Hampton**, **Smith**, **Cleeremans & Destrebecqz**). Rules would be more abstract (**Bailey**) and less sensitive to context than Similarity. Functionally, pure Rules operations on novel instances would be indistinguishable from the same operations on old ones (**Cleeremans & Destrebecqz**). In a Rules system our requirement for internal consistency is typically more stringent than in Similarity knowledge (**Sloman**). The applicability of a Rule is usually restricted, but Similarity comparisons are generally very flexible. Additionally, the R-S proposal has a number of implications that are less intuitive: Rules operations could be supported by many exemplars in the same way as Similarity operations – it's just that for Rules operations these exemplars would be a lot more similar to each other. Rules effects would show some gradedness, even if not as much as for Similarity operations. Could we test an r-s proposal in terms of predictions like the above (cf. **Heit & Hayes**)? Since there is definitional arbitrariness in such a proposal, I suggest that the important issue is to agree on which behaviors should be characterized rules and which similarity. In other words, an r-s proposal is primarily a descriptive tool in psychological theory.

Accordingly, an r-s proposal should be evaluated along the following two dimensions: First, does it provide a principled characterization of processes as rules or similarity with the least number of assumptions, consistently with our intuitions of which processes should be considered rules and which similarity? This issue was discussed in the target article. Second, does it enable interesting implications about rules and similarity behavior? The primary aim of this paper was to demonstrate how the R-S proposal could be integrated with inferential machinery (coherence, goals, and commitment) to predict when Rules would be preferred relative to Similarity. It is suggested that an r-s proposal is successful only to the extent to which it can address these questions and that no more formal investigation of the proposal is possible.

In closing, even though the R-S proposal has been criticized on account of its simplicity, it has been successfully applied to a wide range of empirical data to a reasonable degree of specificity. And the R-S proposal is most certainly not meant to exhaust r-s theorizing. Instead, it can serve as a principled framework through which we can further develop our understanding of these operations (the present utilization of coherence, goals, and commitment is an illustration of this point).

References

Letters "a" and "r" appearing before authors' initials refer to target article and response, respectively.

Aha, D. W. & Goldstone, R. L. (1992) Concept learning and flexible weighting. In: *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, ed. J. K. Kruschke, pp. 534–39. Erlbaum. [aEMP]

Albright, A. & Hayes, B. (2003) Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90:119–61. [rEMP]

Allen, S. W. & Brooks, L. R. (1991) Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General* 120:3–19. [aEMP]

Altmann, G. T. & Dienes, Z. (1999) Rule learning by seven-month-old infants and neural networks. *Science* 284:87. [aEMP]

Anderson, A. K. & Phelps, E. A. (2001) Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature* 411:305–309. [OV]

Anderson, J. R. (1983) *The architecture of cognition*. Harvard University Press. [ABM]

(1990) *The adaptive character of thought*. Erlbaum. [aEMP]

(1993) *Rules of the mind*. Erlbaum. [aEMP]

Anderson, J. R. & Lebiere, C. (2003) The Newell test for a theory of cognition. *Behavioral and Brain Sciences* 26(5):587–640. [FCG]

Anderson, J. R., Reder, L. M. & Lebiere, C. (1996) Working memory: Activation limitations on retrieval. *Cognitive Psychology* 30(3):221–56. [ABM]

https://doi.org/10.1017/S0140525X05470016 Published online by Cambridge University Press

Andruski, J. E., Blumstein, S. E. & Burton, M. W. (1994) The effect of subphonetic differences on lexical access. *Cognition* 52:163–87. [BM]

Arló-Costa, H. (1996) Epistemic conditionals, snakes and stars. In: *Conditionals, from philosophy to computer science. Studies in logic and computation, vol. 5*, ed. L. Fariñas del Cerro, G. Crocco & A. Herzig, pp. 193–239. Oxford University Press. [HA-C]

(1999) Belief revision conditionals: Basic iterated systems. *Annals of Pure and Applied Logic* 96:3–28. [HA-C]

(2001) Bayesian epistemology and epistemic conditionals: On the status of the export-import laws. *Journal of Philosophy* 98(11):555–98. [HA-C]

Arló-Costa, H. & Shapiro, S. (1992) Maps between conditional logic and non-monotonic logic. In: *Principles of knowledge representation and reasoning: Proceedings of the Third International Conference*, ed. B. Nebel, C. Rich & W. Swartout, pp. 553–65. Morgan Kaufmann. [HA-C]

Armstrong, S. L., Gleitman, L. R. & Gleitman, H. (1983) What some concepts might not be. *Cognition* 13:263–308. [aEMP]

Ashby, F. G. & Alfonso-Reese, L. A. (1995) Categorization as probability density estimation. *Journal of Mathematical Psychology* 39:216–33. [aEMP]

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U. & Waldron, E. M. (1998) A neuropsychological theory of multiple systems in category learning. *Psychological Review* 105:442–81. [FGA, EH]

Ashby, F. G. & Ell, S. W. (2002) Single versus multiple systems of category learning: Reply to Nosofsky and Kruschke (2001) *Psychonomic Bulletin and Review* 9:175–80. [FGA]

Ashby, F. G., Ell, S. W. & Waldron, E. M. (2003) Procedural learning in perceptual categorization. *Memory and Cognition* 31:1114–25. [FGA]

Ashby, F. G., Maddox, W. T. & Bohil, C. J. (2002) Observational versus feedback training in rule-based and information-integration category learning. *Memory and Cognition* 30:666–77. [FGA, JAH]

Ashby, F. G. & Perrin, N. A. (1988) Towards a unified theory of similarity and recognition. *Psychological Review* 95:124–50. [aEMP]

Ashby, F. G., Queller, S. & Berretty, P. M. (1999) On the dominance of unidimensional rules in unsupervised categorization. *Perception and Psychophysics* 61:1178–99. [FGA]

Bailey, T. M. (1995) *Nonmetrical constraints on stress*. Doctoral dissertation, University of Minnesota, Minneapolis. UMI. Available at: http://www.cf.ac.uk/psych/ssd/PAPERS/thesis.html [TMB]

Bailey, T. M., Plunkett, K. & Scarpa, E. (1999) A cross-linguistic study in learning prosodic rhythms: Rules, constraints and similarity. *Language and Speech* 42:1–38. [TMB]

Baker, C. L. & McCarthy, J. J. (1981) *The logical problem of language acquisition*. MIT Press. [aEMP]

Barsalou, L. W. (1985) Ideals, central tendency and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory and Cognition* 11:629–54. [aEMP]

(1991) Deriving categories to achieve goals. In: *The psychology of learning and motivation, vol. 27*, ed. G. H. Bower, pp. 1–64. Academic Press. [aEMP]

(1999) Perceptual symbol systems. *Behavioral and Brain Sciences* 22:577–609. [OV]

(2003) Abstraction in perceptual symbol systems. *Philosophical Transactions of the Royal Society of London: Biological Sciences* 358:1177–87. [OV]

Bassok, M., Chase, V. M. & Martin, S. A. (1998) Adding apples and oranges: Semantic constraints on application of formal rules. *Cognitive Psychology* 35(2):99–134. [ABM]

Berko, J. (1958) The child's learning of English morphology. *Word* 14:150–77. [aEMP]

Bisiacchi, P., Denes, G. & Semenza, C. (1976) Semantic fields in aphasia: An experimental investigation on comprehension of the relation of class and property. *Archives Suisse de Neurologie, Neurosurgerie et Psychiatrie* 118:207–13. [JD]

Bock, J. K. (1986) Syntactic persistence in language production. *Cognitive Psychology* 18:355–87. [aEMP]

Bod, R. (1998) *Beyond grammar: An experience-based theory of language*. Center for the Study of Language and Information Publications. [UH]

Boole, G. (1854) *An investigation of the laws of thought*. Dover. [aEMP]

Booth, A. E. & Waxman, S. R. (2002) Word learning is "smart": Evidence that conceptual information affects preschoolers' extension of novel words. *Cognition* 84:B11–B22. [GD]

Braine, M. D. S. (1978) On the relation between the natural logic of reasoning and standard logic. *Psychological Review* 85:1–21. [aEMP]

Braine, M. D. S., O'Brien, D. P., Noveck, I. A., Samuels, M. C., Lea, B. L., Fisch, S. M. & Yang, Y. (1995) Predicting intermediate and multiple conclusions in propositional logic inference problems: Further evidence for a mental logic. *Journal of Experimental Psychology: General* 124:263–92. [aEMP]

Braisby, N., Franks, B. & Hampton, J. (1996) Essentialism, word use and concepts. *Cognition* 59:247–74. [aEMP]

Breedin, S. D., Saffran, E. M. & Coslett, B. (1994) Reversal of the concreteness effect in a patient with semantic dementia. *Cognitive Neuropsychology* 11:617–60. [JD]

Brooks, L. R. & Hannah, S. D. (2000) Relation between perceptual and informational learning of family resemblance structures. Paper presented at 41st Annual Meeting of the Psychonomics Society, New Orleans, LA. [LRB]

(submitted) Instantiated features and the use of "rules." [LRB]

Brooks, R. L. & Vokey, R. J. (1991) Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: Learning, Memory and Cognition* 120:316–23. [LRB, aEMP, PR]

Bruner, J., Goodnow, J. & Austin, G. (1956) *A study of thinking*. Transaction Publishers. [JAH]

Burzio, L. (2002) Missing players: Phonology and the past-tense debate. *Lingua* 112:157–99. [GL]

Byrne, R. M. J. (1989) Suppressing valid inferences with conditionals. *Cognition* 31:61–83. [aEMP]

Calvo, F. & Colunga, E. (2003) The statistical brain: Reply to Marcus' *The Algebraic Mind*. In: *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society*, ed. R. Alterman & D. Kirsh, pp. 210–15. Erlbaum. [FCG]

Campion, J., Lane, R., Brander, G. & Koritsas, E. (1996) A task model to develop principles of decision aiding. In: *Proceedings of the Second International Command and Control Research and Technology Symposium, Market Bosworth, UK, September 24–26, 1996*. [JC]

Caramazza, A. & Mahon, B. Z. (2003) The organization of conceptual knowledge: Evidence from category-specific semantic deficits. *Trends in Cognitive Sciences* 7:354–61. [JD]

Carruthers, P. (2000) *Phenomenal consciousness: A naturalistic theory*. Cambridge University Press. [DED]

Cazden, C. B. (1968) The acquisition of noun and verb inflections. *Child Development* 39:433–48. [aEMP]

Chater, N. (1997) Simplicity and the mind. *The Psychologist* 495–98. [aEMP]

(1999) The search for simplicity: A fundamental cognitive principle? *Quarterly Journal of Experimental Psychology* 52A:273–302. [rEMP]

Chater, N. & Hahn, U. (1997) Representational distortion, similarity and the Universal Law of Generalization. In: *Proceedings of the Similarity and Categorization Workshop*, eds. M. Ramscar, U. Hahn, E. Cambouropolous & H. Pain. 97: 31–36. University of Edinburgh. [aEMP]

Chater, N. & Oaksford, M. (2000) The rational analysis of mind and behavior. *Synthese* 122:93–131. [EH]

Cheng, P. W. & Holyoak, K. J. (1985) Pragmatic reasoning schemas. *Cognitive Psychology* 17:391–416. [aEMP]

Chertkow, H., Bub, D., Deaudon, C. & Whitehead, V. (1997) On the status of object concepts in aphasia. *Brain and Language* 58:203–32. [JD]

Chomsky, N. (1957) *Semantic structures*. Mouton. [aEMP]

(1965) *Aspects of the theory of syntax*. MIT Press. [aEMP]

Chomsky, N. & Miller, G. A. (1958) Finite state languages. *Information and Control* 1:91–112. [aEMP]

Churchland, P. M. (1990) Cognitive activity in artificial grammar neural networks. In: *An invitation to cognitive science, vol. 3: Thinking*, ed. D. N. Osherson & E. E. Smith. MIT Press. [aEMP]

Clahsen, H. (1999) The dual nature of the language faculty: A case study of German inflection. *Behavioral and Brain Sciences* 22:991–1013. [GM]

Clark, A. & Karmiloff-Smith, A. (1993) The cognizer's innards: A philosophical and psychological perspective on the development of thought. *Mind and Language* 8(4):487–519. [AC]

Cleeremans, A. (1997) Principles for implicit learning. In: *How implicit is implicit knowledge?*, ed. D. Berry, pp. 95–234. Oxford University Press. [AC]

Cleeremans, A. & McClelland, J. L. (1991) Learning the structure of event sequences. *Journal of Experimental Psychology: General* 120:235–53. [AC]

Cohen, R., Kelter, S. & Woll, G. (1980) Analytical competence and language impairment in aphasia. *Brain and Language* 10:341–47. [JD]

Cowan, N. (2001) The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24:87–185. [WD]

Crain, S. (1991) Language acquisition in the absence of experience. *Behavioral and Brain Sciences* 14:597–650. [aEMP]

Cronbach, L. J. & Meehl, P. (1955) Construct validation in psychological tests. *Psychological Bulletin* 52:281–382. [DED]

Cross, C. & Nute, D. (1998) Conditional logic. In: *Handbook of philosophical logic: Vol. 3: Extensions of classical logic, second edition*, ed. D. Gabbay & F. Guenthner. Reidel. [HA-C]

Dahan, D., Magnuson, J. S., Tanenhaus, M. K. & Hogan, E. M. (2001) Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes* 16:507–34. [BM]

Davidoff, J. & Roberson, D. (2004) Preserved thematic and impaired taxonomic categorisation: A case study. *Language and Cognitive Processes* 19:137–74. [JD]

Davies, M. (1995) Two notions of implicit rule. In: *Philosophical perspectives, vol.*

https://doi.org/10.1017/S0140525X05470016 Published online by Cambridge University Press

9: *Connectionist and philosophical psychology*, ed. J. Tomberlin. Ridgeview [aEMP]
Dienes, Z. (1992) Connectionist and memory-array models of Artificial Grammar Learning. *Cognitive Science* 16:41–79. [aEMP]
Dienes, Z., Altmann, G. T. & Gao, S. (1999) Mapping across domains without feedback: A neural network model of transfer of implicit knowledge. *Cognitive Science* 23:53–82. [PFD, aEMP]
Diesendruck, G., Gelman, S. A. & Lebowitz, K. (1998) Conceptual and linguistic biases in children's word learning. *Developmental Psychology* 34:823–39. [GD]
Diesendruck, G., Markson, L. & Bloom, P. (2003) Children's reliance on creator's intent in extending names for artifacts. *Psychological Science* 14:164–68. [GD]
Dominey, P. F., Lelekov, T., Ventre-Dominey, J. & Jeannerod, M. (1998) Dissociable processes for learning the surface and abstract structure sensorimotor sequences. *Journal of Cognitive Neuroscience* 10(6):734–51. [PFD]
Driver, J. (1996) Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature* 381:66–68. [OV]
Duch, W., Adamczak, R. & Grabczewski, K. (2001) A new methodology of extraction, optimization and application of crisp and fuzzy logical rules. *IEEE Transactions on Neural Networks* 12:277–306. [WD]
Duch, W. & Grudziński, K. (2001) Prototype based rules – A new way to understand the data. In: *Proceedings of the International Joint Conference on Neural Networks, Washington, D.C., 15–19 July 2001*, ed. K. Marko & P. Werbos, pp. 1858–63. IEEE Press. [WD]
Duch, W., Setiono, R. & Zurada, J. M. (2004) Computational intelligence methods for rule-based data understanding. *Proceedings of the IEEE* 92:771–805. [WD]
Dulany, D. E. (1997) Consciousness in the explicit (deliberative) and implicit (evocative). In: *Scientific approaches to consciousness*, ed. J. Cohen & J. Schooler, pp. 179–212. Erlbaum. [DED]
(1999) Consciousness, connectionism, and intentionality. *Behavioral and Brain Sciences* 22:154–55. [DED]
(2003) Strategies for putting consciousness in its place. *Journal of Consciousness Studies* 10:33–43. [aEMP]
(2004) Higher order representation in a mentalistic metatheory. In: *Higher order thought theories of consciousness*, ed. R. J. Gennaro, pp. 315–38. John Benjamins. [DED]
Dulany, D. E., Carlson, R. A. & Dewey, G. I. (1984) A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General* 113:541–55. [aEMP]
Dummett, M. (1975) Wang's paradox. *Synthese* 30:301–24. [JD]
Elman, J. L. (1996) *Rethinking innateness: A connectionist perspective on development*. MIT Press. [aEMP]
Erickson, M. A. & Kruschke, J. K. (1998) Rules and exemplars in category learning. *Journal of Experimental Psychology: General* 127:107–40. [WD, aEMP]
Estes, W. K. (1994) *Classification and cognition*. Oxford University Press. [EES]
Evans, J. St. B. T. (1972) Interpretation and matching bias in a reasoning task. *British Journal of Psychology* 24:193–99. [aEMP]
(1991) Theories of human reasoning: The fragmented state of the art. *Theory and Psychology* 1:83–105. [aEMP]
Evans, J. St. B. T., Newstead, S. E. & Byrne, R. J. M. (1991) *Human reasoning: The psychology of deduction*. Erlbaum. [aEMP]
Evans, J. St. B. T. & Over, D. E. (1996) *Rationality and reasoning*. Psychology Press. [EH]
Fodor, J. & Pylyshyn, Z. (1988) Connectionism and cognitive architecture: A critique. *Cognition* 28:3–71. [FCG, aEMP]
Franklin, S., Howard, D. & Patterson, K. (1995) Abstract word anomia. *Cognitive Neuropsychology* 12:549–66. [JD]
Freeman, W. J. (2000) *Neurodynamics: An exploration in mesoscopic brain dynamics*. Springer-Verlag. [FCG]
Gabbay, D. M. (1985) Theoretical foundations for non-monotonic reasoning in expert systems. In: *Proceedings of the NATO Advanced Study Institute on Logics and Models of Concurrent Systems*, ed. K. R. Apt, pp. 439–59. Springer-Verlag. [HA-C]
Gallaway, C. & Richards, B. (1994) *Input and interaction in language acquisition*. Cambridge University Press. [aEMP]
Gardner, H. & Zurif, E. B. (1976) Critical reading of words and phrases in aphasia. *Brain and Language* 3:173–90. [JD]
Gati, I. & Tversky, A. (1984) Weighting common and distinctive features in perceptual and conceptual judgments. *Cognitive Psychology* 16:341–470. [ABM]
Gelman, S. A. & Markman, E. M. (1986) Categories and induction in young children. *Cognition* 23:183–209. [GD]
Gelman, S. A. & Wellman, H. M. (1991) Insides and essences: Early understanding of the non-obvious. *Cognition* 38:213–44. [aEMP]
Gentner, D. (1983) Structure-mapping: A theoretical framework for analogy. *Cognitive Science* 7:155–70. [ABM]
(1988) Metaphor as structure mapping: The relational shift. *Child Development* 48:1034–39. [JD]
Gentner, D. & Markman, A. B. (1997) Structural alignment in analogy and similarity. *American Psychologist* 52(1):45–56. [ABM]
Gentner, D. & Medina, J. (1998) Similarity and the development of rules. *Cognition* 65:263–97. [JD, EH, aEMP]
Gigerenzer, G. & Regier, T. (1996) How do we tell an association from a rule? Comment on Sloman (1996) *Psychological Bulletin* 119:23–26. [aEMP]
Gluck, M. A. & Bower, G. H. (1988) From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General* 8:37–50. [DED]
Goel, V., Gold, B., Kapur, S. & Houle, S. (1997) The seats of reason: A localization study of deductive and inductive reasoning using PET (O15) blood flow technique. *NeuroReport* 8:1305–10. [EH]
Gold, E. M. (1967) Language identification in the limit. *Information and Control* 16:447–74. [aEMP]
Goldstein, K. (1948) *Language and language disturbances*. Grune and Stratton. [JD]
Goldstone, R. L. (1994a) The role of similarity in categorization: Providing a groundwork. *Cognition* 52:125–57. [JAH, aEMP]
(1994b) Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General* 123:178–200. [aEMP]
(1995) Effects of categorization on color perception. *Psychological Science* 6:298–304. [aEMP]
(1998) Perceptual learning. *Annual Review of Psychology* 49:585–612. [GL, rEMP]
(1999) Similarity. In: *MIT encyclopaedia of the cognitive sciences*, ed. R. A. Wilson & F. C. Keil, pp. 763–65. MIT Press. [GL, OV]
Goldstone, R. L. & Barsalou, L. W. (1998) Reuniting perception and conception. *Cognition* 65:231–62. [aEMP]
Goldstone, R. L., Medin, D. L. & Gentner, D. (1991) Relational similarity and the non-independence of features in similarity judgments. *Cognitive Psychology* 23:222–64. [LRB]
Goldstone, R. L. & Son, J. (in press) Similarity. In: *Cambridge handbook of thinking and reasoning*, ed. K. Holyoak & R. Morrison. Cambridge University Press. [OV]
Goldstone, R. L., Steyvers, M. & Larimer, K. (1996) Categorical perception of novel dimensions. In: *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*, ed. G. W. Cottrell, pp. 243–48. Erlbaum. [aEMP]
Goodglass, H., Hyde, M. R. & Blumstein, J. (1969) Frequency, picturability and availability of nouns in aphasia. *Cortex* 5:104–19. [JD]
Goodman, N. (1972) Seven strictures on similarity. In: *Problems and projects*, ed. N. Goodman, pp. 437–47. Bobbs-Merrill. [JAH, aEMP]
Gow, D. (2001) Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language* 45:133–39. [BM]
(2002) Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance* 28(1):163–79. [BM]
Gow, D. W. Jr. (2003) Feature parsing: Feature cue mapping in spoken word recognition. *Perception and Psychophysics* 65(4):575–90. [BM]
Gow, D. W. Jr. & Gordon, P. C. (1995) Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance* 21(2):344–59. [BM]
Gow, D. W. & Holcomb, P. (2002) Electrophysiological correlates of place assimilation context effects. Paper presented at the *43rd Annual Meeting of the Psychonomic Society*, November, 2002, Kansas City, MO. [BM]
Gow, D. W., McMurray, B. & Tanenhaus, M. K. (2003) Eye movements reveal the time course of multiple context effects in the perception of assimilated speech. Poster presented at The 44th Annual Meeting of the Psychonomics Society, November, 2003, Vancouver, Canada. [BM]
Griggs, R. A. (1983) The role of problem content in the selection task and in the THOG problem. In: *Thinking and reasoning: Psychological approaches*, ed. J. St. B. T. Evans. Routledge. [aEMP]
Griggs, R. A. & Cox, J. R. (1982) The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology* 73:407–20. [aEMP]
Grossman, M. (1981) A bird is a bird is a bird: Making reference with and without superordinate categories. *Brain and Language* 12:313–31. [JD]
Grossman, M., Smith, E. E., Koenig, P., Glosser, G., DeVita, L., Moore, P. & McMillan, C. (2002) The neural basis for categorization in semantic memory. *Neuroimage* 17:1549–61. [EES]
Guenther, F. H., Husain, F. T., Cohen, M. A. & Shinn-Cunningham, B. G. (1999) Effects of categorization and discrimination training on auditory perceptual space. *Journal of the Acoustical Society of America* 106:2900–12. [GL]
Gupta, P. & Touretzky, D. S. (1994) Connectionist models and linguistic theory: Investigations of stress systems in language. *Cognitive Science* 18:1–50. [TMB]

https://doi.org/10.1017/S0140525X05470016 Published online by Cambridge University Press

Hadley, R. (1993) The "explicit / implicit" distinction. Technical Report CSS-IS TR93–02, Simon Frasier University. [aEMP]
(1999) Connectionism and novel combinations of skills: Implications for cognitive architecture. *Minds and Machines* 9:197–221. [FCG]
Hahn, U. & Chater, N. (1998) Similarity and rules: Distinct? Exhaustive? Empirically distinguishable? *Cognition* 65:197–230. [aEMP, RR, OV, UH]
Hahn, U. & Nakisa, R. C. (2000) German inflection: Single or dual route? *Cognitive Psychology* 41:313–60. [aEMP]
Hahn, U., Prat-Sala, M. & Pothos, E. M. (2002) How similarity affects the ease of rule application. In: *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*, eds. W. D. Gray & C. D. Schunn, Erlbaum. [aEMP]
Halle, M. & Vergnaud, J. R. (1987) *An essay on stress*. MIT Press. [TMB]
Hammer, R. & Diesendruck, G. (2005) The role of dimensional distinctiveness in children's and adults' artifact categorization. *Psychological Science* 16:137–44. [GD]
Hannah, S. D. & Brooks, L. R. (submitted a) Biasing of categorization decisions: Mediation by the perceptual familiarity of features and by learned attention patterns. [LRB]
(submitted b) Feature appearance as a determiner of feature importance in classification. [LRB]
Hare, M., Elman, J. L. & Daugherty, K. G. (1995) Default generalization in connectionist networks. *Language and Cognitive Processes* 10:601–30. [aEMP]
Harman, G. (1999) *Reasoning, meaning, and mind*. Oxford University Press. [EH]
Harnad, S., ed. (1987) *Categorical perception*. Cambridge University Press. [aEMP]
Heit, E. (1997) Knowledge and concept learning. In: *Knowledge, concepts, and categories*, ed. K. Lamberts & D. Shanks, pp. 7–41. Psychology Press. [EH, aEMP]
(1998) A Bayesian analysis of some forms of inductive reasoning. In: *Rational models of cognition*, ed. M. Oaksford & N. Chater, pp. 248–74. Oxford University Press. [EH]
Hempel, C. G. (1952) *Fundamentals of concept formation in the empirical sciences*. University of Chicago Press. [DED]
Henaff Gonon, M. A., Bruckert, R. & Michel, F. (1989) Lexicalization in an anomic patient. *Neuropsychologia* 27:391–407. [JD]
Henle, M. (1962) On the relation between logic and thinking. *Psychological Review* 69:366–78. [aEMP]
Herrnstein, R. J., Vaughan, W. J. R., Mumford, D. B. & Kosslyn, S. M. (1989) Teaching pigeons an abstract relational rule: Insideness. *Perception and Psychophysics* 46:56–64. [aEMP]
Hinton, G. E. (1986) Learning distributed representations of concepts. Paper presented at the 8th Annual Conference of the Cognitive Science Society, Amherst, MA, August 1986. [AC]
Hintzman, D. L. (1986) "Schema abstraction" in a multiple-trace memory model. *Psychological Review* 93:411–28. [arEMP]
Howard, D. & Patterson, K. E. (1992) *The Pyramids and Palm Trees Test*. Thames Valley Test Company. [JD]
Hummel, J. E. & Holyoak, K. J. (1997) Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review* 104(3):427–66. [TMB, ABM]
Inhelder, B. & Piaget, J. (1958) *The growth of logical thinking*. Basic Books. [aEMP]
Isham, C. J. (1989) *Lectures on groups and vector spaces*. World Scientific. [aEMP]
Jackendoff, R. S. (1975) Morphological and semantic regularities in the lexicon. *Language* 51:639–71. [aEMP]
Joanisse, M. F. & Seidenberg, M. S. (1999) Impairments in verb morphology after brain injury: A connectionist model. In: *Proceedings of the National Academy of Sciences USA* 96:7592–97. [rEMP]
Johansen, M. K. & Palmeri, T. J. (2002) Are there representational shifts during category learning? *Cognitive Psychology* 45:482–553. [EH, UH]
Johnson-Laird, P. N. (1993) Précis of *Deduction*. *Behavioral and Brain Sciences* 16:323–80. [aEMP]
(1994) Mental models and probabilistic thinking. *Cognition* 50:189–209. [EH]
Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Legrenzi, M. A. & Caverni, J. P. (1999) Naïve probability: A mental model theory of extensional reasoning. *Psychological Review* 106:62–88. [EH]
Johnstone, T. & Shanks, D. R. (1999) Two mechanisms in implicit grammar learning? Comment on Meulemans and Van der Linden (1997) *Journal of Experimental Psychology: Leaning, Memory, and Cognition* 25:524–31. [aEMP]
Juola, P. & Plunkett, K. (1998) Why double dissociations don't mean much. In: *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*, ed. M. A. Gernsbacher & S. J. Derry, pp. 561–66. Erlbaum. [rEMP]
Kahneman, D. & Tversky, A. (1972) Subjective probability: A judgment of representativeness. *Cognitive Psychology* 3:430–54. [aEMP]
Kaplan, A. S. & Murphy, G. L. (2000) Category learning with minimal prior knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26:829–46. [aEMP]
Katz, J. (1972) *Semantic theory*. Harper & Row. [aEMP]
Keil, F. C. (1989) *Concepts, kinds and cognitive development*. MIT Press. [aEMP]
(1995) The growth of causal understanding of natural kinds. In: *Causal cognition*, ed. D. Sperber, D. Premack & A. Premack. Oxford University Press. [GD]
Keil, F. C., Smith, C., Simons, D. J. & Levin, D. T. (1998) Two dogmas of conceptual empiricism: Implications for hybrid models of the structure of knowledge. *Cognition* 65:103–35. [aEMP]
Kelso, J. A. S. (1995) *Dynamic patterns: The self-organization of brain and behavior*. MIT Press. [FCG]
Kelter, S., Cohen, R., Engel, D., List, G. & Strohner, H. (1976) Aphasic disorders in matching tasks involving conceptual analysis and covert naming. *Cortex* 12:383–94. [JD]
Kim, N. S. & Ahn, W. K. (2002) Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology: General* 131(4):451–76. [ABM]
Knowlton, B. & Squire, L. (1994) The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20:79–91. [aEMP, RR]
(1996) Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22:169–81. [aEMP, RR]
Koepcke, K. M. (1988) Schemes in German plural formation. *Lingua* 74:303–35. [aEMP]
Kolodner, J. L. (1992) An introduction to Case-Based reasoning. *Artificial Intelligence Review* 6:3–34. [aEMP]
Komatsu, L. K. (1992) Recent views of conceptual structure. *Psychological Bulletin* 112:500–26. [aEMP]
Kruschke, J. K. (1992) ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review* 99:22–44. [FGA, EH, aEMP]
Ladefoged, P. & Broadbent, D. E. (1957) Information conveyed by vowels. *Journal of the Acoustical Society of America* 29:98–104. [GL]
Lamberts, K. (2000) Information-accumulation theory of speeded categorization. *Psychological Review* 107:227–60. [EH]
Lamberts, K. & Shanks, D., eds. (1997) *Knowledge, concepts, and categories*. MIT Press. [OV]
Leland, J. (1994) Generalized similarity judgments: An alternative explanation for choice anomalies. *Journal of Risk and Uncertainty* 9:151–72. [HA-C]
Levi, I. (1980) *The enterprise of knowledge*. MIT Press. [HA-C]
(1986) The paradoxes of Allais and Ellsberg. *Economics and Philosophy* 2:23–56. [HA-C]
(1996) *For the sake of the argument: Ramsey test conditionals, inductive inference, and non-monotonic reasoning*. Cambridge University Press. [HA-C]
Lewis, D. (1973) *Counterfactuals*. Blackwell. [HA-C]
Liberman, A. M., Harris, K. S., Hoffman, H. S. & Griffith, B. C. (1957) The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54(5):358–68. [BM]
Liberman, A. M., Harris, K. S., Kinney, J. & Lane, H. (1961) The discrimination of relative onset-time of the components of certain speech and non-speech patterns. *Journal of Experimental Psychology* 61:379. [BM]
Livingston, K. R., Andrews, J. K. & Harnad, S. (1998) Categorical perception effects induced by category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24:732–53. [aEMP]
Macaluso, E., Frith, C. & Driver, J. (2000) Modulation of human visual cortex by crossmodal spatial attention. *Science* 289:1206–1208. [OV]
Mackintosh, N. J. (1983) *Conditioning and associative learning*. Oxford University Press. [aEMP]
MacWhinney, B. (1993) The (il)logical problem of language acquisition. In: *Proceedings of the 15th Annual Conference of the Cognitive Science Society*, ed. M. C. Polsen, Erlbaum. [aEMP]
(2001) The competition model: The input, the context, and the brain. In: *Cognition and second language instruction*, ed. P. Robinson, pp. 69–90. Cambridge University Press. [PR]
Maddox, W. T., Ashby, F. G. & Bohil, C. J. (2003) Delayed feedback effects on rule-based and information-integration category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29:650–62. [FGA]
Maddox, W. T., Ashby, F. G., Ing, A. D. & Pickering, A. D. (2004a) Disrupting feedback processing interferes with rule-based but not information-integration category learning. *Memory and Cognition* 32:582–91. [FGA]
Maddox, W. T., Bohil, C. J. & Ing, A. D. (2004b) Evidence for a procedural learning-based system in perceptual category learning. *Psychonomic Bulletin and Review* 11:945–52. [FGA]

https://doi.org/10.1017/S0140525X05470016 Published online by Cambridge University Press

Malt, B. C. (1990) Features and beliefs in the mental representations of categories. *Journal of Memory and Language* 29:289–315. [aEMP]
Marcus, G. F. (2001) *The algebraic mind: Integrating connectionism and cognitive science*. MIT Press. [PFD, GM, aEMP]
Marcus, G. F. & Berent, I. (2003) Are there limits to statistical learning? *Science* 300:53–54. [FCG]
Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R. & Pinker, S. (1995) German inflection: The exception that proves the rule. *Cognitive Psychology* 29:189–256. [aEMP]
Marcus, G. F., Vijayan, S., Bandi Rao, S. & Vishton, P. M. (1999) Rule learning by seven-month-old infants. *Science* 283:77–80. [AC, FCG, aEMP]
Markman, A. B. (2001) Structural alignment, similarity and the internal structure of category representations. In: *Similarity and categorization*, ed. U. Hahn & M. Ramscar, pp. 190–30. Oxford University Press. [JD]
Markman, A. B. & Ross, B. H. (2003) Category use and category learning. *Psychological Bulletin* 129:592–615. [EH]
Marr, D. (1982) *Vision: A computational investigation into the human representation and processing of visual information*. W. H. Freeman. [arEMP]
McClelland, L. J. & Rumelhart, E. D. (1986) Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General* 114:159–88. [aEMP]
McClelland, J. L., St. John, M. & Taraban, R. (1989) Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes* 4:287–335. [GL]
McGee, V. (1985) A counterexample to modus ponens. *Journal of Philosophy* 82:462–71. [HA-C]
McKinley, S. C. & Nosofsky, R. M. (1995) Investigations of exemplar and decision bound models in large, ill-defined category structures. *Journal of Experimental Psychology: Human Perception and Performance* 21:128–48. [aEMP]
McMurray, B., Aslin, R., Tanenhaus, M., Spivey, M. & Subik, D. (in preparation). Two "B" or not 2 /b/: Categorical perception in lexical and nonlexical tasks. [BM]
McMurray, B., Tanenhaus, M. & Aslin, R. (2002) Gradient effects of within-category phonetic variation on lexical access. *Cognition* 86(2):B33–42. [BM]
McMurray, B., Tanenhaus, M., Aslin, R. & Spivey, M. (2003) Probabilistic constraint satisfaction at the lexical/phonetic interface: Evidence for gradient effects of within-category VOT on lexical access. *Journal of Psycholinguistic Research* 32(1):77–97. [BM]
Medin, D. L., Goldstone, R. L. & Gentner, D. (1993) Respects for similarity. *Psychological Review* 100(2):254–78. [ABM]
Medin, D. L. & Ortony, A. (1989) Psychological essentialism. In: *Similarity and analogical reasoning*, ed. S. Vosniadou & A. Ortony. Cambridge University Press. [aEMP]
Medin, D. L. & Ross, B. H. (1989) The specific character of abstract thought: Categorization, problem solving and induction. In: *Advances in the psychology of human intelligence, vol. 5,* ed. R. S. Sternberg, pp. 189–223. Erlbaum. [ABM, aEMP]
Medin, D. L. & Schaffer, M. M. (1978) Context theory of classification learning. *Psychological Review* 85(3):207–38. [aEMP, EES]
Medin, D. L. & Smith, E. E. (1981) Strategies and classification learning. *Journal of Experimental Psychology: Human Learning and Memory* 7:241–53. [aEMP]
Medin, D. L., Wattenmaker, W. D. & Hampson, S. E. (1987) Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology* 19:242–79. [EH]
Meulemans, T. & van der Linden, M. (1997) Associative chunk strength in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23:1007–28. [aEMP, RR]
Miozzo, M. (2003) On the processing of regular and irregular forms of verbs and nouns: Evidence from neuropsychology. *Cognition* 87:101–27. [rEMP]
Mitchell, T. (1997) *Machine learning*. McGraw Hill. [WD]
Murphy, G. L. & Medin, D. L. (1985) The role of theories in conceptual coherence. *Psychological Review* 92:289–316. [arEMP]
Newell, A. (1973) You can't play 20 questions with nature and win: Projective comments on the papers in this symposium. In: *Visual information processing*, ed. W. G. Chase, pp. 283–308. Academic Press. [JGW]
(1990) *Unified theories of cognition*. Harvard University Press. [rEMP]
Nijenhuis, J. T. & van der Flier, H. (2002) The correlation of g with attentional and perceptual-motor ability tests. *Journal of Personality and Individual Differences* 33:287–97. [rEMP]
Nosofsky, R. M. (1986) Attention, similarity and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115(1):39–57. [DED, EES]
(1988) Similarity, frequency, and category representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14:54–65. [aEMP]
(1989) Further tests of an exemplar-similarity approach to relating identification and categorization. *Journal of Experimental Psychology: Perception and Psychophysics* 45:279–90. [aEMP]
(1990) Relations between exemplar-similarity and likelihood models of classification. *Journal of Mathematical Psychology* 34:393–418. [aEMP]
(1991) Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance* 17:3–27. [FCG, arEMP]
Nosofsky, R. M., Clark, S. E. & Shin, H. J. (1989) Rules and exemplars in categorization, identification, and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2:282–304. [JAH, aEMP]
Nosofsky, R. M. & Johansen, M. K. (2000) Exemplar-based accounts of "multiple-system" phenomena in perceptual categorization. *Psychonomic Bulletin and Review* 7:375–402. [EH]
Nosofsky, R. B. & Zaki, S. R. (1998) Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science* 9:247–55. [aEMP]
Nygaard, L. C. & Pisoni, D. B. (1998) Talker-specific learning in speech perception. *Perception and Psychophysics* 60:355–76. [GL]
Oaksford, M. & Chater, N. (1994) A rational analysis of the selection task as optimal data selection. *Psychological Review* 101:608–31. [arEMP, SS]
Osherson, D. N. (1990) Judgment. In: *Thinking: An invitation to cognitive science*, ed. D. N. Osherson & E. E. Smith. MIT Press. [aEMP]
Osherson, D. N., Perani, D., Cappa, S., Schnur, T., Grassi, F. & Fazio, F. (1998) Distinct brain loci in deductive versus probabilistic reasoning. *Neuropsychologia* 36:369–76. [EH]
Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A. & Shafir, E. (1990) Category-based induction. *Psychological Review* 97(2):185–200. [EH, ABM, aEMP]
Pacton, S., Perruchet, P., Fayol, M. & Cleeremans, A. (2001) Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General* 130(3):401–26. [AC]
Paris, J. B. (1994) *The uncertain reasoner's companion: A mathematical perspective*. Cambridge University Press. [aEMP]
Parsons, L. M. & Osherson, D. N. (2001) New evidence for distinct right and left brain systems for deductive versus probabilistic reasoning. *Cerebral Cortex* 11:954–65. [EH]
Patalano, A. L., Smith, E. E., Jonides, J. & Koeppe, R. A. (2002) PET evidence for multiple strategies of categorization. *Cognitive, Affective and Behavioural Neuroscience* 1(4):360–70. [EES]
Pearce, J. M. (1987) A model for stimulus generalization in Pavlovian conditioning. *Psychological Review* 94:61–73. [aEMP]
Perruchet, P. & Pacteau, C. (1990) Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General* 119:264–75. [TMB, aEMP]
Perruchet, P. & Vinter, A. (2002) The self-organizing consciousness. *Behavioral and Brain Sciences* 25(3):297–329. [DED]
Pickering, M. J. & Branigan, H. P. (1999) Syntactic priming in language production. *Trends in Cognitive Sciences* 3:136–42. [aEMP]
Pinker, S. (1979) Formal models of language learning. *Cognition* 7:217–83. [aEMP]
(1994) *The language instinct*. Morrow. [aEMP]
(1999) *Words and rules: The ingredients of language, 1st edition*. Basic Books. [GM, EES]
Pinker, S. & Prince, A. (1988) On language processing and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28:73–193. [aEMP]
Pinker, S. & Ullman, M. (2002) The past and future of the past tense. *Trends in Cognitive Sciences* 6:456–63. [FCG]
Plunkett, K. & Marchman, V. A. (1991) U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition* 38:43–102. [WD, aEMP]
(1993) From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition* 48(1):21–69. [aEMP]
Plunkett, K & Nakisa, R. C. (1997) A connectionist model of the Arabic plural system. *Language and Cognitive Processes* 12:807–36. [aEMP]
Pollard, P. (1982) Human reasoning: Some possible effects of availability. *Cognition* 12:65–96. [aEMP]
Posner, M. I. & Keele, S. W. (1968) On the genesis of abstract ideas. *Journal of Experimental Psychology* 77:353–63. [aEMP]
Pothos, E. M. & Bailey, T. M. (2000) The importance of similarity in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26:847–62. [aEMP]
Pothos, E. M. & Chater, N. (2002) A simplicity principle in unsupervised human categorization. *Cognitive Science* 26:303–43. [arEMP]
Pothos, E. M. & Hahn, U. (2000) So concepts aren't definitions, but do they have necessary *or* sufficient features? *British Journal of Psychology* 91:439–50. [aEMP]
Prince, A. & Smolensky, P. (1993) Optimality theory: Constraint interaction in

https://doi.org/10.1017/S0140525X05470016 Published online by Cambridge University Press

generative grammar. Technical Report No. 2. Rutgers University Center for Cognitive Science. [TMB]
Putnam, H. (1975) The meaning of "meaning." In: *Mind, language and reality: Philosophical papers, vol. 2*, ed. H. Putnam. Cambridge University Press. [aEMP]
Ramscar, M. (2002) The role of meaning in inflection: Why the past tense does not require a rule. *Cognitive Psychology* 45:45–94. [FCG, GL, aEMP]
Reber, A. R., Allen, R. & Regan, S. (1985) Syntactical learning and judgment, still unconscious and still abstract: Comment on Dulany, Carlson, and Dewey. *Journal of Experimental Psychology: General* 114:17–24. [rEMP]
Reber, A. S. (1967) Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior* 6:317–27. [aEMP, RR]
(1989) Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General* 118:219–35. [aEMP, PR]
Reber, A. S. & Allen, R. (1978) Analogic and abstraction strategies in synthetic grammar learning: A functional interpretation. *Cognition* 6:189–221. [aEMP]
Redington, F. M. & Chater, N. (1996) Transfer in artificial grammar learning: Methodological issues and theoretical implications. *Journal of Experimental Psychology: General* 125:123–38. [aEMP]
Reed, S. K. (1972) Pattern recognition and categorization. *Cognitive Psychology* 3:382–407. [aEMP]
Regehr, G. & Brooks, L. R. (1995) Category organization in free classification: The organizing effect of an array of stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* 21:347–63. [UH]
Rips, L. J. (1983) Cognitive processes in propositional reasoning. *Psychological Review* 90:38–71. [aEMP]
(1989) Similarity, typicality, and categorization. In: *Similarity and analogical reasoning*, ed. S. Vosniadou & A. Ortony, pp. 21–59. Cambridge University Press. [DED, aEMP, EES]
(1994) *The psychology of proof: Deductive reasoning in human thinking*. MIT Press. [aEMP]
(2001) Necessity and natural categories. *Psychological Bulletin* 127:827–52. [aEMP]
Rips, L. J. & Collins, A. (1993) Categories and resemblance. *Journal of Experimental Psychology: General* 122:468–86. [aEMP]
Roberson, D., Davidoff, J. & Braisby, N. (1999) Similarity and categorisation: Neuropsychological evidence for a dissociation in explicit categorisation tasks. *Cognition* 71:1–42. [JD]
Robinson, P. (1996) *Consciousness, rules, and instructed second language acquisition*. Lang. [PR]
(2002) Effects of individual differences in intelligence, aptitude, and working memory on adult incidental SLA: A replication and extension of Reber, Walkenfield and Hernstadt, 1991. In: *Individual differences and instructed language learning*, ed. P. Robinson, pp. 211–66. John Benjamins. [PR]
(2005) Cognitive abilities, chunking, and frequency effects in artificial grammar and incidental L2 learning: Replications of Reber, Wakenfeld, and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance to SLA. *Studies in Second Language Acquisition* 27(2):235–68. [PR]
Rosch, E. & Mervis, C. B. (1975) Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7:573–605. [aEMP]
Rubinstein, A. (1988) Similarity and decision making under risk (Is there a utility theory resolution of the Allais paradox?). *Journal of Economic Theory* 46:145–53. [HA-C]
(1997) *Modeling bounded rationality*. MIT Press. [HA-C]
Rumelhart, D. E. & McClelland, J. L. (1986) On learning the past tenses of English verbs: Implicit rules or parallel distributed processing? In: *Parallel distributed processing: Explorations in the microstructure of cognition. vol. 2: Psychological and biological models*, ed. J. L. McClelland, D. E. Rumelhart & the PDP Research Group. MIT Press. [aEMP, UH]
Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996) Statistical learning by 8-month old infants. *Science* 274:1926–28. [aEMP]
Schalkoff, R. (1992) *Pattern recognition. Statistical, structural and neural approaches*. Wiley. [WD]
Schank, R. C. (1982) *Dynamic memory: A theory of learning in people and computers*. Cambridge University Press. [aEMP]
Schneider, W., Chein, J. & McHugo, M. (2003) A model of automatic/controlled processing in the brain. Paper presented at the 44th Annual Meeting of the Psychonomic Society, Vancouver, Canada, November 6–9, 2003. [DED]
Schroyens, W., Schaeken, W. & Handley, S. (2003) In search of counter examples: Deductive rationality in human reasoning. *The Quarterly Journal of Experimental Psychology* 56A:1129–45. [rEMP]
Schyns, P. G., Goldstone, R. L. & Thibaut, J. (1998) Development of features in object concepts. *Behavioral and Brain Sciences* 21:1–54. [EH, GL]
Schyns, P. G. & Oliva, A. (1999) Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition* 69:243–65. [aEMP]
Searle, J. (1980) Rules and causation. *Behavioral and Brain Sciences* 3:1–61. [aEMP]
(1992) *The rediscovery of the mind*. MIT Press. [AC]
Seidenberg, M. S. & Elman, J. L. (1999) Networks are not "hidden rules." *Trends in Cognitive Sciences* 3:288–89. [FCG]
Seidenberg, M. S., MacDonald, M. C. & Saffran, J. R. (2003) Are there limits to statistical learning? – Response. *Science* 300:54. [FCG]
Semenza, C., Bisiacchi, P. & Romani, L. (1992) Naming disorders and semantic representations. *Journal of Psycholinguistic Research* 21:349–64. [JD]
Semenza, C., Denes, G., Lucchese, D. & Bisiacchi, P. (1980) Selective deficit of conceptual structures in aphasia: Class versus thematic relations. *Brain and Language* 10:243–48. [JD]
Servan-Schreiber, E. & Anderson, J. R. (1990) Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16:592–608. [TMB]
Shanks, D. R. & Darby, R. J. (1998) Feature- and rule-based generalization in human associative learning. *Journal of Experimental Psychology: Animal Behavior Processes* 24:405–15. [aEMP]
Shanks, D. R., Johnstone, T. & Kinder, A. (2002) Modularity and artificial grammar learning. In: *Implicit learning: An empirical, philosophical and computational consensus in the making*, ed. R. M. French & A. Cleeremans, pp. 93–120. Psychology Press. [RR]
Shanks, D. R., Johnstone, T. & Staggs, L. (1997) Abstraction processes in artificial grammar learning. *Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology* 50(1):216–52. [AC]
Shanks, D. R. & St. John, M. F. (1994) Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences* 17:367–447. [DED]
Shepard, R. N. (1987) Toward a universal law of generalization for psychological science. *Science* 237:1317–23. [aEMP]
Singley, M. K. & Anderson, J. R. (1989) *The transfer of cognitive skill.* Harvard University Press. [ABM]
Skinner, B. F. (1936) Conditioning and extinction and their relation to drive. *Journal of General Psychology* 14:296–317. [aEMP]
(1957) *Verbal behavior*. Appleton. [JC]
Sloman, S. A. (1993) Feature-based induction. *Cognitive Psychology* 25:231–80. [EH]
(1996) The empirical case for two systems of reasoning. *Psychological Bulletin* 119:3–22. [EH, aEMP]
Sloman, S. A. & Rips, L. J. (1998) Similarity as an explanatory construct. *Cognition* 65:87–101. [aEMP]
Smith, E. E., Langston, C. & Nisbett, R. E. (1992) The case for rules in reasoning. *Cognitive Science* 16:1–40. [aEMP, EES]
Smith, E. E. & Medin, D. L. (1981) *Categories and concepts.* Harvard University Press. [EES]
Smith, E. E. & Osherson, D. N. (1989) Similarity and decision making. In: *Similarity and analogical reasoning*, ed. S. Vosniadou & A. Ortony, pp. 60–71. Cambridge University Press. [HA-C]
Smith, E. E., Patalano, A. L. & Jonides, J. (1998) Alternative strategies of categorization. *Cognition* 65:167–96. [EH, aEMP, RR]
Smith, E. E. & Sloman, S. A. (1994) Similarity- vs. rule-based categorization. *Memory and Cognition* 22:377–86. [aEMP]
Smith, L. B. (1989) A model of perceptual classification in children and adults. *Psychological Review* 96:125–44. [EH]
Smith, L. B., Jones, S. S. & Landau, B. (1996) Naming in young children: A dumb attentional mechanism. *Cognition* 60:143–71. [GD]
Spalding, T. L. & Murphy, G. L. (1996) Effects of background knowledge on category construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22:525–38. [EH]
Spencer, J. P. & Thelen, E., eds. (2003) Special issue: Connectionist and dynamic systems approaches to development. *Developmental Science* 4(4). [FCG]
Spohn, W. (1988) A general non-probabilistic theory of inductive inference. In: *Causation in decision, belief change and statistics*, ed. W. Harper & B. Skyrms, pp. 105–34. Reidel. [HA-C]
Stanovich, K. E. (1999) *Who is rational? Studies of individual differences in reasoning*. Erlbaum. [EH]
Sternberg, S. (1966) High-speed scanning in human memory. *Science* 153:652–54. [FGA]
Tesar, B. (1997) An iterative strategy for learning metrical stress in optimality theory. In: *The Proceedings of the 21st Annual Boston University Conference on Language Development*, ed. E. Hughes, M. Hughes & A. Greenhill, pp. 615–26. Cascadilla Press. [TMB]
Thelen, E., Schöner, G., Scheier, C. & Smith, L. (2001) The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences* 24(1):1–86. [FCG]
Thelen, E. & Smith, L. (1994) *A dynamic systems approach to the development of perception and action*. MIT Press. [FCG]
Tversky, A. (1977) Features of similarity. *Psychological Review* 84(4):327–52. [ABM, aEMP, OV]
Tversky, A. & Kahneman, D. (1983) Availability: A heuristic for judging the frequency and probability. *Cognitive Psychology* 5:207–232. [aEMP]

https://doi.org/10.1017/S0140525X05470016 Published online by Cambridge University Press

Ullman, M. T. (2001) The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition* 4(2):105–22. [EES]

Vignolo, L. A. (1999) Disorders of conceptual thinking in aphasia. In: *Handbook of clinical and experimental neuropsychology*, ed. G. Denes & L. Pizzamiglio, pp. 273–88. Psychology Press. [JD]

Vilarroya, O. (2002) *The dissolution of mind*. Rodopi (VIBS). [OV]

Vokey, J. R. & Brooks, L. R. (1992) Salience of item knowledge in learning artificial grammar. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20:328–44. [aEMP, PR]

Vroomen, J. & de Gelder, B. (2000) Sound enhances visual perception: Cross-modal effects of auditory organization on visual perception. *Journal of Experimental Psychology: Human Perception and Performance* 26:1583–90. [OV]

Wagner, U., Gais, S., Haider, H., Verleger, R. & Born, J. (2004) Sleep inspires insight. *Nature* 427(6972):352–55. [AC]

Waldron, E. M. & Ashby, F. G. (2001) The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychonomic Bulletin and Review* 8:168–76. [FGA]

Warrington, E. K. (1975) The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology* 27:635–57. [JD]

Wason, P. C. (1960) On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology* 12:129–40. [JAH, aEMP]

(1966) Reasoning. In: *New horizons in psychology, vol. 1*, pp. 135–51. Penguin. [SS]

Wason, P. C. & Johnson-Laird, P. N. (1972) *The psychology of reasoning: Structure and content*. Harvard University Press. [aEMP]

Wasserman, E. A. & Miller, R. R. (1997) What's elementary about associative learning? *Annual Review of Psychology* 48:573–607. [aEMP]

Whittlesea, B. W. A. & Dorken, M. D. (1993) Incidentally, things in general are particularly determined: An episodic-processing account of implicit learning. *Journal of Experimental Psychology: General* 122:227–48. [RR]

Willingham, D. B., Wells, L. A., Farrell, J. M. & Stemwedel, M. E. (2000) Implicit motor sequence learning is represented in response locations. *Memory and Cognition* 28:366–75. [FGA]

Wills, S. & Mackintosh, N. J. (1998) Peak shift on an artificial dimension. *Quarterly Journal of Experimental Psychology* 51B:1–31. [aEMP]

Wisniewski, E. J. (1995) Prior knowledge and functionally relevant features in concept learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21:449–68. [aEMP]

Wisniewski, E. J. & Medin, D. L. (1994) On the interaction of theory and data in concept learning. *Cognitive Science* 18:221–81. [EH]

Wittgenstein, L. (1953/1998) *Philosophical investigations*. Blackwell. [aEMP]

Wolff, J. G. (1999) Probabilistic reasoning as information compression by multiple alignment, unification and search: An introduction and overview. *Journal of Universal Computer Science* 5(7):418–62. URL: http://arxiv.org/abs/cs.AI/0307010. [JGW]

(2000) Syntax, parsing and production of natural language in a framework of information compression by multiple alignment, unification and search. *Journal of Universal Computer Science* 6(8):781–829. URL: http://arxiv.org/abs/cs.AI/0307014. [JGW]

(2003a) Information compression by multiple alignment, unification and search as a unifying principle in computing and cognition. *Artificial Intelligence Review* 19(3):193–230. URL: http://arxiv.org/abs/cs.AI/0307025. [JGW]

(2003b) Unsupervised grammar induction in a framework of information compression by multiple alignment, unification and search. In: *Proceedings of the workshop and tutorial on learning context-free grammars*, eds. C. de la Higuera, P. Adriaans, M. van Zaanen & J. Oncina. ECML/PKDD 2003, Cavtat-Dubrovnik, Croata. URL: http://arxiv.org/abs/cs.AI/0311045. [JGW]

Wright, C. (1975) On the coherence of vague predicates. *Synthese* 30:325–65. [JD]

Zurif, E., Caramazza, A., Meyerson, R. & Galvin, J. (1974) Semantic feature representation for normal and aphasic language. *Brain and Language* 1:167–87. [JD]

https://doi.org/10.1017/S0140525X05470016 Published online by Cambridge University Press