

Cross-modal information integration in category learning

J. David Smith · Jennifer J. R. Johnston ·
Robert D. Musgrave · Alexandria C. Zakrzewski ·
Joseph Boomer · Barbara A. Church · F. Gregory Ashby

Published online: 27 March 2014
© Psychonomic Society, Inc. 2014

Abstract An influential theoretical perspective describes an implicit category-learning system that associates regions of perceptual space with response outputs by integrating information preattentively and predecisionally across multiple stimulus dimensions. In this study, we tested whether this kind of implicit, information-integration category learning is possible across stimulus dimensions lying in different sensory modalities. Humans learned categories composed of conjoint visual–auditory category exemplars comprising a visual component (rectangles varying in the density of contained lit pixels) and an auditory component (in Exp. 1, auditory sequences varying in duration; in Exp. 2, pure tones varying in pitch). The categories had either a one-dimensional, rule-based solution or a two-dimensional, information-integration solution. Humans could solve the information-integration category tasks by integrating information across two stimulus modalities. The results demonstrated an important cross-modal form of sensory integration in the service of category learning, and they advance the field’s knowledge about the sensory organization of systems for categorization.

Keywords Category learning · Implicit cognition · Cross-modal perception · Sensory integration · Basal ganglia

Categorization is an essential cognitive function—an ancient information-processing capability that remains crucial for survival. For this reason, categorization has received sustained

research attention (Ashby & Maddox, 2010; Brooks, 1978; Feldman, 2000; Homa, Sterling, & Trepel, 1981; Knowlton & Squire, 1993; Murphy, 2002; Nosofsky, 1987; Posner & Keele, 1968; Rosch & Mervis, 1975; Smith et al., 2014; Smith, Redford, & Haas, 2008; Wasserman, Kiedinger, & Bhatt, 1988).

A lasting issue is whether multiple systems are necessary to accommodate the diverse demands of categorization. (A similar issue concerns the possibility of multiple memory systems.) The possibility of multiple category-learning systems is strengthened by categorization’s long natural history spanning vertebrate evolution. Humans’ cognitive system might contain ancestral category-learning capacities in older brain regions and recent primate or hominid capacities.

An emerging consensus does grant humans multiple category-learning systems (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby & Ell, 2001; Erickson & Kruschke, 1998; Homa et al., 1981; Maddox & Ashby, 2004; Rosseel, 2002; Smith, Berg, et al., 2012; Smith & Minda, 1998). This consensus reflects many findings. For example, different processes dominate categorization when the categorization rule is easy or difficult to verbalize (Ashby & Maddox, 2010), when categories are large or small (Blair & Homa, 2003; Homa et al., 1981; Minda & Smith, 2001), and when category training is minimal or extensive (Cook & Smith, 2006; Reed, 1978; Smith, Chapman, & Redford, 2010; Wasserman et al., 1988). Some question multiple-system interpretations of some results (Nosofsky & Johansen, 2000), believing that all categorization phenomena can be explained using a unitary, exemplar-based process (Newell, Dunn, & Kalish, 2010; Nosofsky, Stanton, & Zaki, 2005; but see Smith et al., 2014). Our research informs this multiple-systems debate without depending on any particular theoretical position.

The consensus for multiple systems is strongest in cognitive neuroscience literature that distinguishes implicit and explicit category-learning systems (Ashby et al., 1998; Ashby & Ell, 2001; Ashby & Valentin, 2005; Smith, Beran, Crossley, Boomer,

J. D. Smith (✉) · J. J. R. Johnston · A. C. Zakrzewski · J. Boomer ·
B. A. Church
Department of Psychology, University of Buffalo, The State
University of New York, 346 Park Hall, Buffalo, NY 14260, USA
e-mail: psysmith@buffalo.edu

R. D. Musgrave · F. G. Ashby
Department of Psychological & Brain Sciences,
University of California, Santa Barbara, Santa Barbara, CA, USA

& Ashby, 2010; Smith, Crossley, et al., 2012). The implicit system encodes stimuli through broad attentional processes that encompass multiple stimulus features. It learns slowly, linking behavioral responses to whole (unanalyzed) category exemplars. Its behavioral solutions are unavailable to awareness or declarative cognition. The explicit system encodes stimuli using narrow attentional processes to test hypotheses about informative stimulus features. Working memory and executive attention support these hypothesis tests. Explicit category learning is available to awareness and declarative cognition.

Support for these categorization systems comes from rule-based (RB) and information-integration (II) tasks (Figs. 1a and b). Each category exemplar in these tasks is a conjoint stimulus combining values from perceptual dimensions X and Y .

In Fig. 1a, the horizontal category boundary means that only Y -axis variation carries valid category information. Low and high dimension Y values define the Category A and B members. The participant must discover this dimensional rule from the presentation of single items using the feedback following each response. This is an explicit task, because the solution can be discovered through reasoning and verbalized.

In Fig. 1b, the diagonal category boundary means that both dimensions carry partially valid information. In this II task, one-dimensional hypotheses are nonoptimal. A vertical or horizontal boundary will not adequately partition the categories. To perform optimally, one must combine or integrate information across the stimulus dimensions. Humans often accomplish this integration implicitly and nondeclaratively.

The dissociation between implicit and explicit systems of categorization is grounded empirically (see the review in

Maddox & Ashby, 2004). For example, II category learning is selectively impaired when reinforcement is delayed, even for just a few seconds (Maddox, Ashby, & Bohil, 2003; Maddox & Ing, 2005), when learning is unsupervised (Ashby, Queller, & Berretty, 1999), and when category knowledge is imparted observationally, not through response reinforcement (Ashby, Maddox, & Bohil, 2002). II learning is supported by a time-locked sequence of events (perception, response selection, and reinforcement). RB category learning is robust to reinforcement delays, unsupervised learning conditions, and so forth. It is supported by hypotheses actively held in working memory, and thus is selectively impaired when working memory resources are diverted (Waldron & Ashby, 2001). The RB–II dissociation is also grounded in the study of neuropsychological patient populations (Ashby & Ennis, 2006).

However, theoretical and empirical issues remain. One issue is the literature's restricted stimulus materials (i.e., mainly sine-wave gratings or Gabor patches—DeValois & DeValois, 1988), a restriction ensuring that new stimuli do not bring methodological problems. It is essential that the field broaden the range of stimulus dimensions to which its theoretical proposals apply. The present research contributes to this broadening.

A second issue is the focus on visual dimensions. It is unknown whether the area's theoretical proposals apply to the auditory modality, for example. Another contribution of the present research is that we included auditory stimulus materials.

A third issue concerns limits on the information-integration capacity of categorization. It seems it could be adaptive for an organism to solve category problems by integrating information across sensory modalities. For example, both bitterness

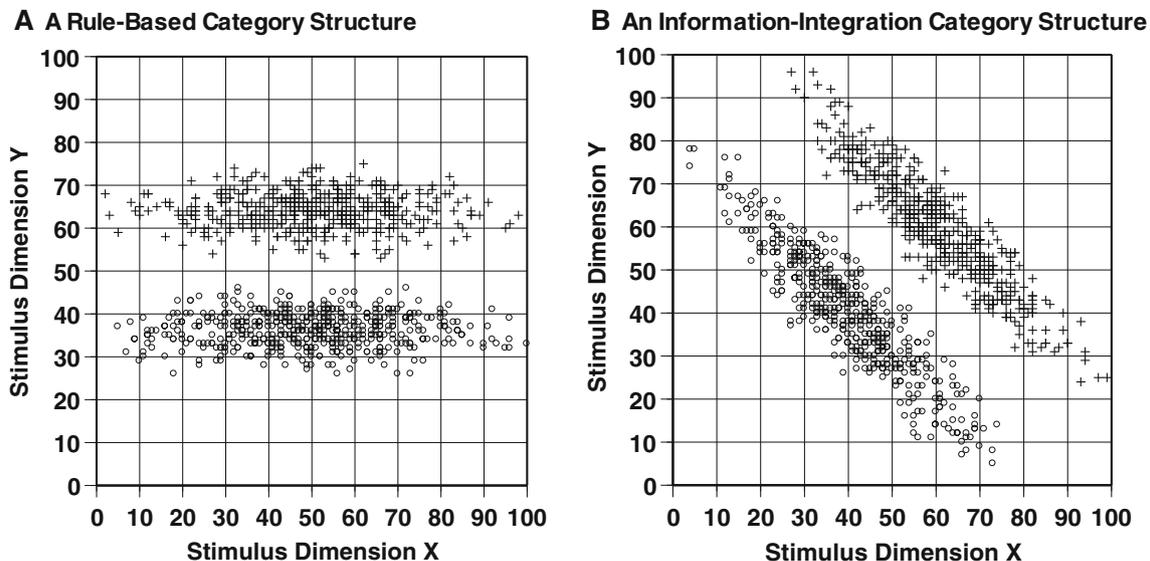


Fig. 1 **a** A rule-based category structure. The open-circle and plus-sign symbols, respectively, indicate Category A and Category B stimuli. The stimuli are depicted within an abstract space, with each dimension's having 101 levels. In the rule-based task, only variation in dimension Y (the auditory dimension, in the present experiments) carries diagnostic category information, so optimal performance would be governed by a

one-dimensional rule (e.g., low- vs. high-pitch auditory stimuli). **b** An information-integration category structure, depicted in the same way. Both dimensions Y and X (the density of illuminated pixels in a box on the screen) carry useful but insufficient category information; thus, information from both dimensions should be integrated into category decisions

and discoloration might betray a food toxin, but if both cues were faint at threshold level, it might require the integration of the cues to trigger a suprathreshold sensory warning. Almost no research has evaluated whether humans can integrate category information across modalities in the service of categorization.

A fourth issue concerns the cognitive level at which humans accomplish sensory integration during categorization. They might do so implicitly, out of awareness, at an early, procedural stage of information processing. Or they might do so explicitly, in awareness, as they made separate decisions about the information provided by each dimension before reaching a category decision. The latter form of integration might represent an explicit cognitive evaluation that could be verbalized as a conjunctive rule. We will consider both kinds of cross-modal integration here.

These issues motivated our study, along with the general concern that scant research has explored cross-modal category learning from any theoretical perspective. One exception was a pioneering study by Maddox, Ing, and Lauritzen (2006). They gave participants a cross-modal II task combining visual spatial frequency (Gabor patches) and auditory frequency. They found some evidence for cross-modal integration. However, their II participants only performed at 59 % correct in their final block of testing and through the last half of the experiment. This is very low performance for II tasks of this kind, though the Maddox et al. (2006) II task was a high-discriminability task that could have allowed high performance. Even the participants fit best by an information-integration model were still only 63 % correct overall and 71 % correct in their last block. Accordingly, Maddox et al. (2006) emphasized the difficulty of cross-modal integration, the possible deficit in integration, and the possibility that the deficit leads participants to adopt rule-based strategies instead. They explained the deficit by pointing to the relative absence of cross-modal neurons or interneurons in the striatum.

Maddox et al. (2006) pointed out that one study cannot fully answer the question of cross-modal integration. The need remains to evaluate broadly whether humans robustly integrate information across modalities during category learning. This was our empirical goal. Our investigation also potentially advances understanding of the function of the basal ganglia in category learning. It is hardly known whether these brain structures—crucial to II category learning—can integrate perceptual input from multiple sensory modalities. The studies could also inform formal models of II category learning, like the COVIS model (Ashby et al., 1998), that have left open whether cross-modal information integration in category learning occurs.

General method

Participants

The participants were 93 and 102 undergraduates (Exps. 1 and 2, respectively) from the University at Buffalo, representative

of the university's psychology participant group, who participated in fulfillment of course requirements. Participants had apparently normal or corrected-to-normal vision, and were placed randomly into RB vertical (RBV, visual dimension relevant), RB horizontal (RBH, auditory dimension relevant), II minor diagonal (II_m), or II major diagonal (II_M) category tasks on the basis of their sequential participant number.

In Experiment 1, participants were dropped from the analysis for not completing 400 trials during the 90-min experimental session (three, five, five, and five participants, respectively, in the four tasks), for showing worse performance in their last 100 trials than in their first 100 trials (zero, zero, five, and three participants in the four tasks), or for never performing above chance in their task (two, one, zero, and zero participants). The final data set included 16 participants in each task. In Experiment 2, participants were dropped from the analyses for not completing 300 trials during the 55-min experimental session (zero, one, one, and zero participants) or for showing worse performance in their last 100 trials than in their first 100 trials (one, zero, two, and four participants). The final data set included 21 participants in each task. These exclusions were like those used in recent RB–II studies (e.g., Smith et al., 2014). They served the goal to observe most clearly the capacity for cross-modal integration in category learning, a goal that noncompleting and nonlearning participants could not advance.

Stimuli

Category exemplars were conjoint visual-auditory stimuli with values on dimensions in different modalities. In Experiments 1 and 2, the visual stimulus was a 200 × 100 pixel unframed box containing randomly placed yellow pixels presented on a black background in the top-center of the computer screen. One hundred and one levels of pixel density could be presented. Each level's pixel count was given by pixels = round(850 × 1.018^{Level}). Pixel density varied from 850 lit pixels (Level 0, 4.2 % illumination of the 20,000 pixels within the 200 × 100 pixel box) to 5,061 lit pixels (Level 100, 25.3 % illumination). The density continuum runs along the *X* dimension in Fig. 1. The visual stimulus components were viewed on a 17-in. screen with an 800 × 600 pixel resolution from a distance of about 24 in.

In Experiment 1, the auditory stimulus comprised three pulsed 1000-Hz tones of the same duration, with pauses among them of the same duration. One hundred and one levels of duration could be presented. Each level's duration was given by the formula (100 + 4*Level)/1,000. For Level 0, the duration was 0.1 s, and the total auditory stimulus (three sounds, two silences) lasted 0.5 s. For Level 100, duration was 0.5 s, and the total auditory stimulus lasted 2.5 s. Sounds were produced using the SOUND procedure within Turbo Pascal 7.0. The continuum of auditory series of slowing pace and increasing duration runs along the *Y* dimension in Fig. 1. On

each trial, the auditory component began immediately following the illumination of the visual component that remained on the screen during the auditory component.

In Experiment 2, the auditory dimension was the frequency of a pure tone. Each level's frequency in Hertz was given by the formula $220 * 2^{(\text{Level} * 10) / (1,200)}$. For Levels 0 and 100, respectively, the pitches were 220 and 392 Hz. The auditory stimulus component was selected from a series of 100 .wav files of eight-bit sine-wave pure tones (mono, 11-kHz sampling rate) created using the wavwrite() function in MATLAB. Volume was set to a comfortable, audible level. Each step along the continuum was ten musical cents, or 0.1 half steps. The frequency continuum spanned 1,000 cents, ten half steps, or a minor 7th interval. Thus we arranged not to flirt with octave effects near 1,200 cents. The tone began at trial onset and terminated either at the participant's response or after 10 s. So that the participant experienced visual and auditory stimulus components from the same spatial locus, the auditory component was presented from a speaker placed directly behind the monitor facing toward the monitor's center. This prevented the direct sight of the speaker and gave the impression that the tone emanated from the pixel box.

Category exemplars were created using Ashby and Gott's (1988) randomization technique. Categories were defined by bivariate normal distributions along the stimulus dimensions, within which each stimulus dimension ranged along a normalized 0–100 scale. Each stimulus was created by drawing a random sample (x, y) from the Category A or Category B distributions that will be specified momentarily. To control for statistical outliers, the random sample was not presented if its Mahalanobis distance from the mean (e.g., Fukunaga, 1972) was greater than 3.0. Instead, another random sample was drawn until the distance criterion was met.

Category structures

Table 1 gives the population distributions that governed the selection of the Category A and B exemplars. These category structures were instantiated using the abstract Level 0–100 scales. Then, as each category exemplar was selected as an x, y coordinate pair in that space, these abstract levels were transformed into concrete conjoint visual and auditory stimuli using the formulae already described. Each participant received his or her own sample of 400 or 300 category exemplars appropriate to the category task that they had been assigned. Through this technique, we granted the RB–II paradigm ecological validity. Individual organisms always experience their own unique statistical sample of a category's probability distribution. The alternative technique—that of giving all participants the same physical stimuli—does not correspond to nature. Moreover, this approach only gives an illusion of control. In fact, organisms always misperceive every stimulus with substantial perceptual error, and therefore

Table 1 Distributional characteristics for the four category tasks

Task	Category	Mean X	Mean Y	Var X	Var Y	Covar XY
<i>RBH</i>	A	50.00	35.86	355.55	16.33	0
	B	50.00	64.14	355.55	16.33	0
<i>RBV</i>	A	35.86	50.00	16.33	355.55	0
	B	64.14	50.00	16.33	355.55	0
<i>IIm</i>	A	40.00	40.00	185.94	185.94	−169.61
	B	60.00	60.00	185.94	185.94	−169.61
<i>IIM</i>	A	40.00	60.00	185.94	185.94	169.6
	B	60.00	40.00	185.94	185.94	169.61

RB, rule-based; *II*, information integration; *H*, horizontal, with the auditory dimension relevant; *V*, vertical, with the visual dimension relevant; *m*, minor diagonal; *M*, major diagonal.

each participant nevertheless experiences his or her own unique sample of subjective stimuli. The formal modeling in this article was always done individually, incorporating each participant's statistical sample of the category task.

Categorization trials

Each trial consisted of one conjoint visual–auditory category exemplar. The response icons (A and B) were located on the screen's lower left and lower right. Participants responded by pressing the “S” or “L” keyboard keys, spatially positioned to correspond to the A or B on the screen, and labeled as “A” or “B.” In Experiment 1, the auditory component of the category exemplar ended when it ended. The visual component of the category exemplar disappeared following a response. On correct responses, the point total incremented, and a 1-s, computer-generated reward whoop sounded. Participants saw a green text banner (Correct +1) centered on the screen along with their accumulated points in white text. On incorrect responses, the point total decremented, in combination with a 1-s, computer-generated penalty sound and an 8-s timeout. Humans saw a red text banner (Incorrect −1) centered on the screen along with their accumulated points in white text. These messages and scorecards appeared as the response icons disappeared following an A or B category response. Experiment 2's categorization trials were nearly identical, except that all auditory feedback signals were eliminated to purify the acoustic environment.

Instructions

Participants were told that each trial would combine a pixel box and a sound, that trials would equally often belong to Category A or B, that they should respond “A” or “B,” and that they would need to guess at first, but with practice and learning would become able to categorize accurately. They were told that they would gain or lose points for correct or incorrect answers, and that incorrect answers would bring 8-s

timeouts. It was pointed out that errors would cost points and lengthen their sessions. Finally, to increase motivation, they were reminded that a \$10 prize would go to the participants who earned points most efficiently.

Formal modeling

To describe the decisional strategies of individual participants, we fit decision bound models (Maddox & Ashby, 1993) to the data from each participant's last 100 trials (when their strategy had matured). These models and our formal procedures are described now.

The *one-dimensional rule-based model* assumes that the participant sets a decision criterion along either the visual or auditory dimension. Modeling specifies the vertical or horizontal line through stimulus space that best partitions the participant's Category A and Category B responses. This model has two parameters: a criterion value on the relevant dimension and a perceptual noise variance. These models fit performances in which participants selectively attend to one dimension and establish a criterion rule along that dimension.

The *conjunctive model* assumes a strategy of the type: "Respond 'A' if the visual level is high enough and the auditory level is high enough, otherwise respond 'B'." This dual criterion carves out a quadrant in the stimulus space reserved for Category A responding. Accordingly, several versions of the conjunctive model can be fit that rotate that response region through different quadrants in stimulus space. These conjunctive models have three parameters: one for the single decision criterion placed along each stimulus dimension, and a perceptual noise variance. When they fit well, these models express that participants achieved a cross-modal integration that occurred after dimensional analysis and selective attention, and after separate criterial decisions on the two dimensions. These postdecisional conjunctions may occur as part of an explicit cognitive evaluation that might be verbalizable and could be characterized as a conjunctive rule.

The *general-linear-classifier (GLC) model* specifies the (nonvertical, nonhorizontal) line through the stimulus space that would best partition the participant's Category A and Category B responses. This model has three parameters: the slope and intercept of the linear decision bound and a perceptual noise variance. When they fit well, these models express that participants achieved another theoretically important kind of information integration across dimensions. This integration occurs before and possibly without dimensional analysis and selective attention, and without separate criterial decisions on the dimensions. This predecisional information integration probably occurs implicitly and may be

based on procedural learning within the striatal category-learning system.

The procedures for selecting the best-fitting model were as follows. Parameters for the models were estimated using the method of maximum likelihood. That is, modeling evaluated which model would, with maximum likelihood, have created the distribution within the stimulus space of Category A and B responses that the participant actually produced. Then the model with the smallest *Bayesian information criterion* (BIC; Schwarz, 1978) was selected as the best-fitting model: $BIC = r \ln N - 2 \ln L$, where r is the number of free parameters, N is the sample size, and L is the likelihood of the model given the data.

Results

Experiment 1

Accuracy-based analyses

The average proportions correct are shown in Fig. 2 for successive 20-trial blocks in the RB and II tasks. The proportion-correct data were analyzed using a two-way analysis of variance (ANOVA) with Trial Block (1–20) as a within-subjects factor and Task Type (RB, II) as a between-subjects factor. The analysis revealed a significant main effect of task, $F(1, 62) = 88.39$, $p < .001$, $\eta_p^2 = .588$. Participants

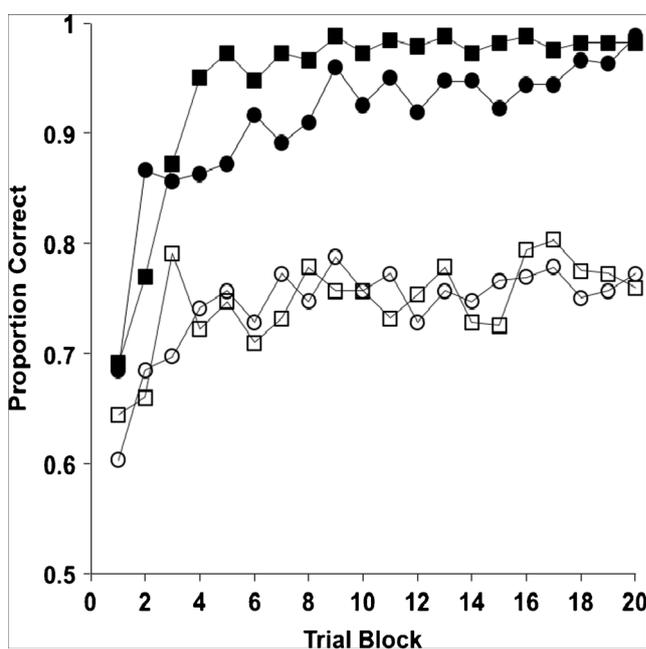


Fig. 2 Proportions of correct responses in each 20-trial block for 16 humans performing 400 trials of a horizontal (auditory) rule-based task (filled circles), a vertical (visual) rule-based task (filled squares), a minor-diagonal information-integration task (open circles), and a major-diagonal information-integration task (open squares)

were .928 and .744 correct overall on the RB and II tasks, respectively. In particular, they were .945 and .911 correct, respectively, on the RBV (visual) and RBH (auditory) tasks, and .743 and .746 correct on the IIm and IIM tasks. The analysis also revealed a significant effect of trial block, $F(19, 1178) = 14.38, p < .001, \eta_p^2 = .188$, with participants showing substantial learning from first block (.655) to last block (.875). The interaction between trial block and task type was weakly present, $F(19, 1178) = 1.91, p = .01, \eta_p^2 = .030$. The reason for this (see Fig. 2) is that RB learning was rapid, so that mainly the RB and II tasks produced parallel (i.e., noninteracting) learning curves. This suggests that stimulus dimensions in different modalities are highly separable in Garner's (1974) sense, making unidimensional category rules in cross-modal RB tasks easy to learn and psychologically insistent. These results replicated the traditional finding of a difference in the speeds of acquisition between RB and II tasks, and a difference in the terminal performance level.

Model-based analyses

The formal modeling let us ask whether participants adopted appropriate decision strategies, whether they integrated dimensional information cross-modally in the II tasks, and whether they likely accomplished this integration explicitly following separate dimensional-criterion tests, or implicitly and procedurally. We modeled the final 100 trials for each participant.

The decision bounds for the RB tasks were organized along the midline of the stimulus space's X (visual) dimension (RBV task, Fig. 3a) or Y (auditory) dimension (RBH task, Fig. 3b). All of the RB participants' data were well fit by a vertical or a horizontal decision bound. The average BIC score was 28.08, indicating strong fits to the data. The RB model was able to account for 97 % of responses, on average, in each participant's last 100 trials. The average percentage correct achieved by these 32 participants across their last 100 trials was 97 %. Participants analyzed these cross-modal stimuli easily and sharply attended

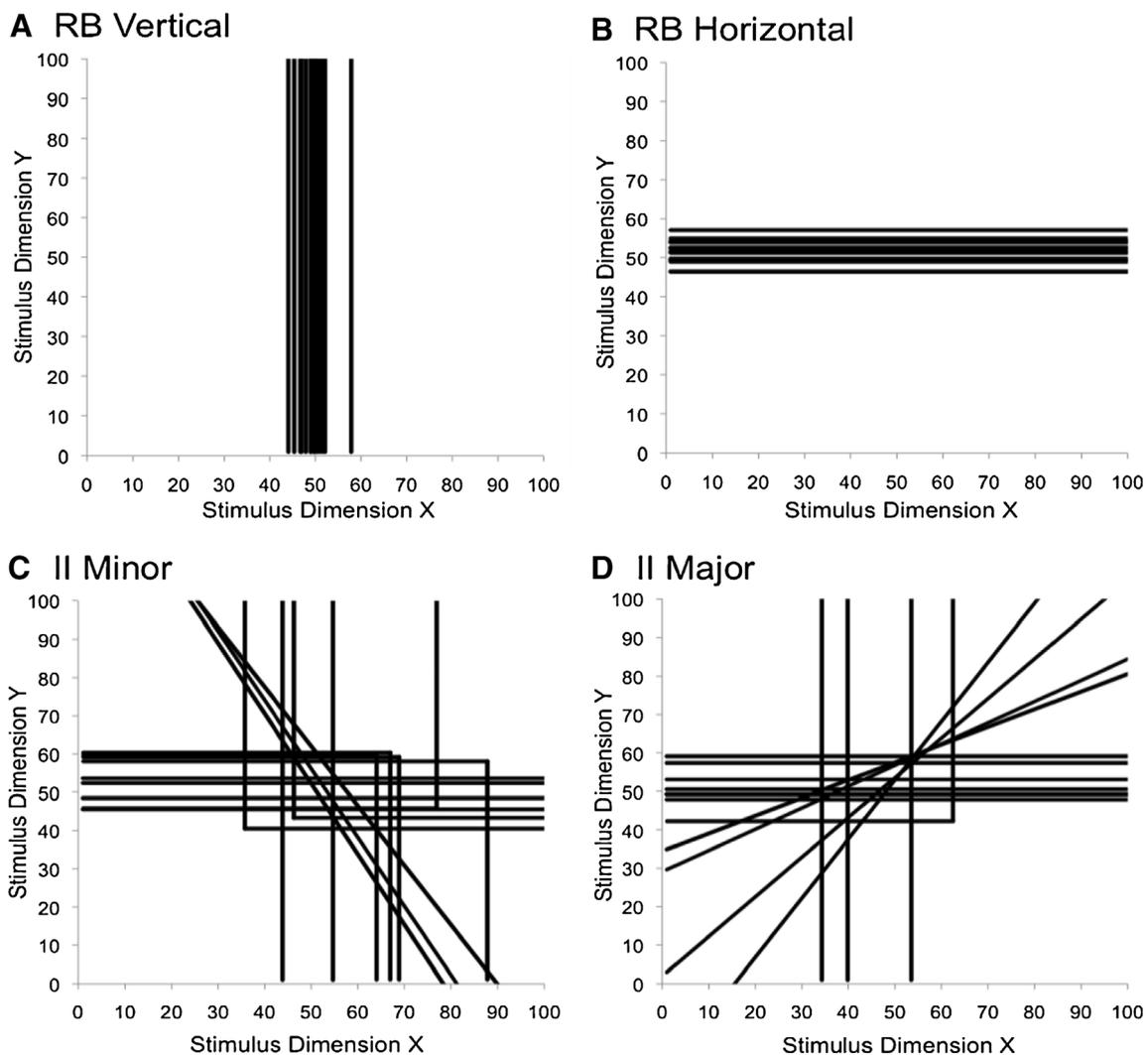


Fig. 3 Decision bounds found in fitting the last 100 responses of 16 participants in the vertical (visual) rule-based task (a), the horizontal (auditory) rule-based task (b), the minor-diagonal information-integration task (c), and the major-diagonal information-integration task (d)

Table 2 Performance and modeling results for the information-integration (II) tasks from Experiments 1 and 2

Exp.	Task	Rule-Based Model				Conjunctive Model				Information-Integration Model			
		% of Sample	Av BIC	% Resp Fit	% Resp Correct	% of Sample	Av BIC	% Resp Fit	% Resp Correct	% of Sample	Av BIC	% Resp Fit	% Resp Correct
E1	IIm	38	86.18	82	70	44	91.10	73	78	19	82.66	86	84
E1	IIM	62	76.02	85	73	6	51.89	64	94	25	73.30	88	87
E2	IIm	38	103.13	77	68	14	84.46	60	76	29	82.72	87	87
E2	IIM	33	76.98	85	73	38	77.48	48	76	19	92.71	84	80

Av BIC, average Bayesian information criterion; % Resp Fit, percentage of responses fit; % Resp correct, percentage of correct responses

to either the visual or auditory component separately and accurately applied their dimensional categorization rule.

The decision bounds for the II tasks were organized more variably, as one can see in the best-fitting decisional boundaries for the IIm task (Fig. 3c) and the IIM task (Fig. 3d). Table 2 (first two rows) summarizes the results of modeling from the II tasks of Experiment 1. The data are broken into participant groups who were best fit by the rule-based model, the conjunctive model, or the information-integration model. Shown are the percentages of the total sample that were best fit by each model, the average BIC scores for those fits, the average percentage of responses for which the model successfully accounted, and the percent correct achieved by that participant group over their last 100 trials. Where the three % sample measures across each row of the table do not sum to 100 %, it is because rare participants were best fit by a random-guessing model, so that a decision bound could not be estimated.

Table 2 shows that the data from 50 % of Experiment 1’s II participants were best fit by an RB boundary. As is common in tasks of this kind, some II participants applied a one-dimensional rule, despite the task’s II organization. Humans have fairly strong rule-based preferences in categorization tasks. These nonoptimal strategies produced poor performance—an average proportion correct of 71 % in the last 100 trials.

The data from 47 % of the II-task participants were best fit by a decision boundary that indicated some form of integration across the visual and auditory stimulus aspects. The integration strategies produced strong performance—an average proportion correct of 83 % over the last 100 trials, compared to 71 % for the participants in the II task who opted less adaptively for a one-dimensional, rule-based strategy. It is an extremely interesting fact—with theoretical, applied, and educational implications—that explicit, declarative cognition frequently intrudes during II performance to narrow attention and systematically disrupt and impair performance.

Finally, we will comment on the kinds of cross-modal integration shown in Experiment 1. Eight of the 15 integrators were best fit by a conjunctive model, suggesting that they devoted a quadrant defined by a dimension *X* and a dimension *Y* criterion as their Category A response region. These participants performed at 80 % accuracy, on average. Seven of the 15 integrators

were best fit by the GLC model, indicating that their decision bound was a true diagonal through the stimulus space. These participants performed at 86 % accuracy, on average—they were easily the strongest II performers. Thus, we found evidence for two kinds of cross-modal integration in our data. Some integration probably occurred postdecisionally, after separate judgments along each dimension. Some integration also probably occurred predecisionally, within the implicit category-learning system.

Experiment 2

Accuracy-based analyses

The aggregate proportions correct for all participants are shown in Fig. 4 for successive 20-trial blocks in each task.

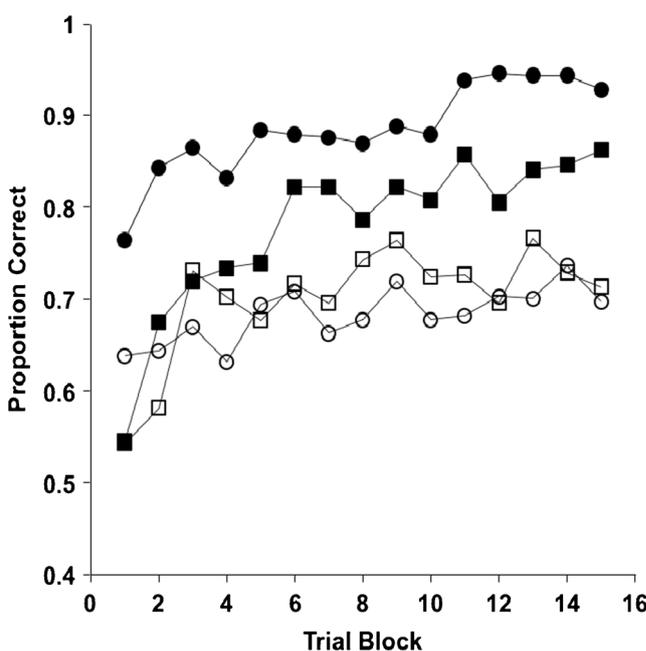


Fig. 4 Proportions of correct responses in each 20-trial block for 21 humans who performed 300 trials of a horizontal (auditory) rule-based task (filled circles), a vertical (visual) rule-based task (filled squares), a minor-diagonal information-integration task (open circles), and a major-diagonal information-integration task (open squares)

The data were analyzed using a two-way ANOVA with Trial Block (1–15) as a within-subjects factor and Task Type (RB, II) as a between-subjects factor. The analysis showed a significant main effect of task, $F(1, 82) = 22.39, p < .001, \eta_p^2 = .214$, with participants being .832 and .691 correct overall on the RB and II tasks, respectively. In particular, they were .778 and .885 correct on the RBV (visual) and RBH (auditory) tasks, and .682 and .700 correct on the IIm and IIM tasks. The analysis also revealed a significant effect of trial block, $F(14, 1148) = 18.12, p < .001, \eta_p^2 = .181$, with participants showing substantial learning from their first block (.623) to their last block (.8). The interaction between trial block and task type was weakly present, as in Experiment 1, $F(14, 1148) = 2.23, p = .006, \eta_p^2 = .027$. Again, we replicated the traditional finding of a difference in learning speeds between RB and II tasks and a difference in terminal performance.

Model-based analyses

The decision bounds for the RB tasks were organized along the midline of the stimulus space's X (visual) dimension (RBV task, Fig. 5a) or Y (auditory) dimension (RBH task, Fig. 5b). The data from the last 100 trials of 38 participants were well fit by a vertical or horizontal decision bound. (The data from four “guessing” participants are not included in the description that follows.) The average BIC score was 39.78, indicating strong fits to the data. The RB model was able to account for 93 % of responses, on average, in each participant's last 100 trials. The average percentage correct achieved by these 38 participants during their last 100 trials was 94 %. Here, too, participants were able to attend selectively to either the visual or the auditory component and to apply accurately a dimensional categorization rule based on that component.

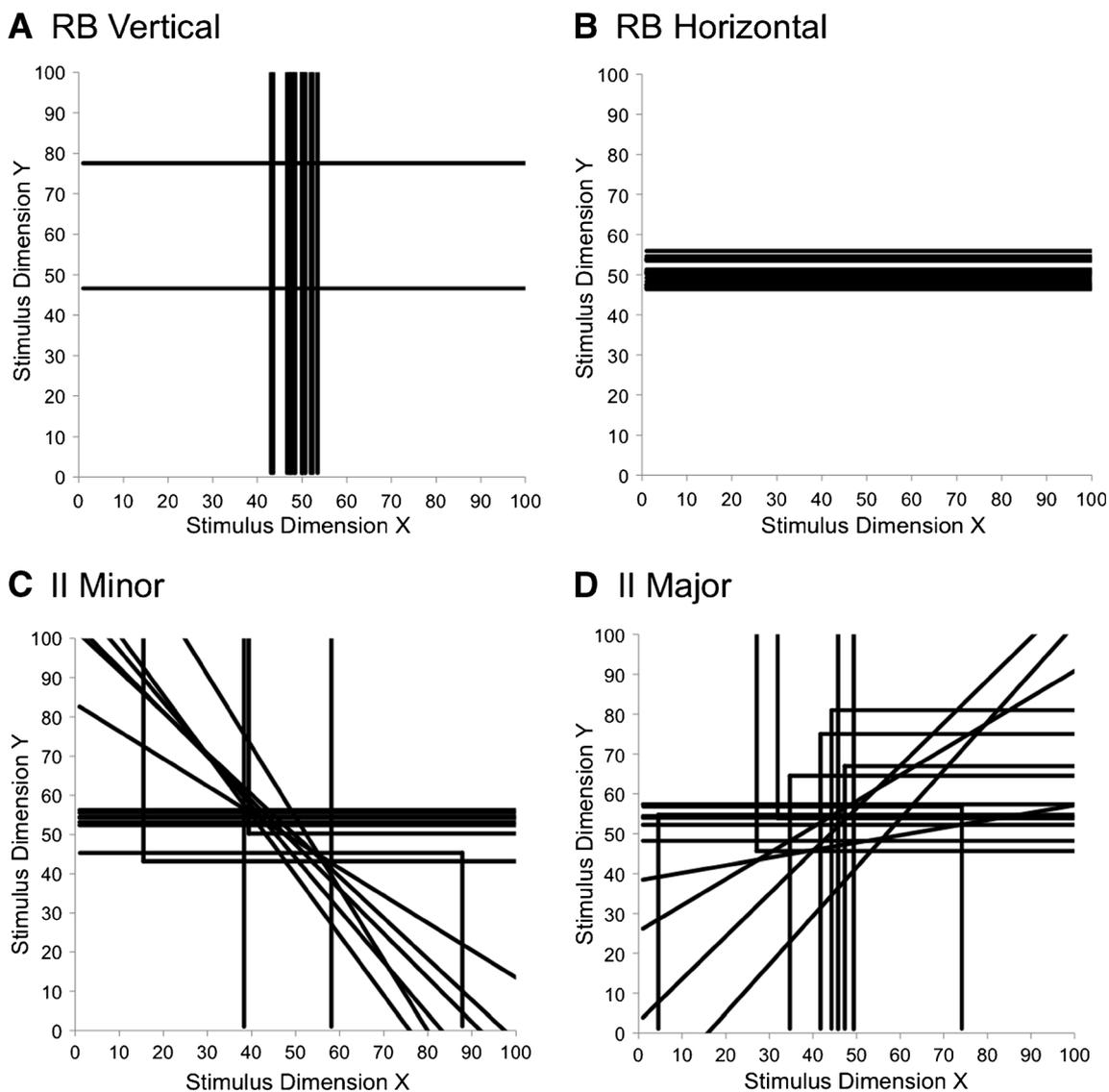


Fig. 5 Decision bounds found in fitting the last 100 responses of 21 participants in the vertical (visual) rule-based task (a), the horizontal (auditory) rule-based task (b), the minor-diagonal information-integration task (c), and the major-diagonal information-integration task (d)

The decision bounds for the II tasks were organized more variably, as one can see in the best-fitting decisional boundaries for the IIm task (Fig. 5c) and the IIM task (Fig. 5d). Table 2 summarizes the results of modeling from the II tasks of Experiment 2. The structure of Table 2 was described previously.

Table 2 shows that the data from 36 % of the II-task participants in Experiment 2 were best fit by a unidimensional, RB boundary. Once again, many participants expressed their analytic preference by applying one-dimensional rules, despite the task's II organization. This nonoptimal approach produced poor performance—an average proportion correct of 70 %, even over the last 100 trials.

The data from 50 % of the II-task participants were best fit by a decision boundary that indicated some form of integration across the visual and auditory stimulus components. This integration strategy produced strong performance—an average percentage correct of 80 % over the last 100 trials, as compared to 70 % for the participants in the II task who opted less adaptively for a one-dimensional, rule-based strategy.

Finally, we will comment on the kinds of cross-modal integration shown in Experiment 2. Eleven of the 21 integrators were best fit by a conjunctive model, suggesting that they devoted a quadrant defined by a dimension X and a dimension Y criterion as their Category A response region. These participants performed at 76 % accuracy, on average. Ten of the 21 integrators were best fit by the GLC model, indicating that their decision bound was a true diagonal through the stimulus space. These participants performed at 84 % accuracy, on average; as in Experiment 1, these participants were easily the strongest II performers. Thus, we found evidence for two kinds of cross-modal integration in our data. Some integration probably occurred postdecisionally, after separate judgments along each dimension, and some probably occurred predecisionally, within the implicit category-learning system. All of these results mirror those of Experiment 1.

Discussion

In two experiments, we gave participants multidimensional category tasks composed of visual and auditory stimulus dimensions. These tasks embodied a strong test of the capacity of humans' category-learning system to integrate information across incommensurate stimulus dimensions into an overall category decision.

The results showed that the RB categorization utility is highly competitive in humans' overall categorization system. It sometimes controls responding adventitiously and nonoptimally, as in our II tasks, wherein one-dimensional rules were unjustified and damaged performance (see also Jitsumori, 1994). Indeed, the data showed high rates of RB defection in the II task. In Smith et al. (2010) and Paul,

Boomer, Smith, and Ashby (2011), these defection rates were near 15 % with purely visual stimuli. Here, the rate was substantially higher with multimodal stimuli (see also Maddox et al., 2006). Multimodal stimuli are easily analyzed into dimensional components, making RB strategies attractive. These stimuli occupy a position of extreme separability along Garner's (1974) integrality-to-separability continuum, explaining why only about half of the participants integrated cross-modally. Individual differences in preference for dimensional rules may determine participants' strategic choices, so that different performance standards might be at work. Some participants may suffice with a 70 %-accurate RB strategy in an II task, whereas others may persist toward an 86 %-accurate II strategy. Diverse individual-difference questions are raised by the present results.

Given their analyzability, multimodal stimuli pose the ultimate challenge to information integration during category learning. Therefore, it is a crucial demonstration that nearly half of the II participants showed adaptive information integration in their categorization. Their integrative strategies produced strong performance—averaging 81 % correct over the last 100 trials. Modeling suggests that these integrative strategies were a mix of explicit, postdecisional and implicit, procedural strategies. The latter cases could provide the strongest confirmation yet that humans can sometimes implicitly integrate the stimulus information provided by multimodal stimulus dimensions into an adaptive categorization principle.

Our two experiments were complementary. Experiment 1 incorporated as an auditory dimension: the duration or pace of successive auditory events. However, this dimension might not have been processed as a purely auditory dimension, but rather as an abstract (amodal) signal of rapidity. This would make participants' information integration in Experiment 1 not fully cross-modal. Therefore, Experiment 2 considered information integration between visual density and auditory pitch, both primary perceptual dimensions. There, participants achieved equally strong levels of cross-modal integration.

Our experiments also complemented those in Maddox et al. (2006), the other empirical study on this topic. A reservation about that study is that participants performed poorly in the II task (about 60 % correct overall). Indeed, even the participants modeled as being information integrators in that study, who presumably brought the optimal strategy to bear on the II task, were only 71 % correct in their last block, as compared to the corresponding final performance levels that we observed of 80 %, 84 %, 86 %, and 87 % (see Table 2). Observing these performances, Maddox et al. (2006) may have estimated too conservatively the capacity for cross-modal integration during category learning. In fact, showing caution, they did emphasize participants' deficits in cross-modal integration, never asserting that they had found strong evidence for such integration. But, in combination with the results in Maddox et al. (2006), our results allow one to more confidently assert the

existence of predecisional, cross-modal integration in category learning.

Meaningful differences distinguish the paradigm used here from that of Maddox et al. (2006). The II task in Maddox et al. (2006) had high discriminability, with a large separation between the Category A and Category B centers in perceptual space (Maddox et al., 2006, p. 1177, Fig. 1a). Their category structure corresponded closely to the low-overlap (i.e., high-separation) structure studied by Ell and Ashby (2006). The latter researchers showed (p. 1018, Fig. 4) that these category structures increase the psychological pull of dimensional analysis, so that RB strategies become dominant. This probably occurs because the RB strategy provides a reasonable level of correct responding, relative to the II strategy. This pull toward RB strategies can even act in some cases to reduce performance in an II task (Maddox, Bohil, & Ing, 2004). Our II tasks had lower discriminability, with less Category A and B separation. They were exactly like those that Ell and Ashby described as *medium-low overlap categories* that should make the RB strategy less competitive and attractive. In fact, Table 2 shows that the II strategies in our tasks produced substantially higher levels of correct responding than did the RB strategies. We may have found more positive evidence of information integration in the present study for this reason. Maddox (personal communication, Dec. 2013) shares this interpretation, and we are grateful for his suggesting it to us. Nonetheless, Maddox et al. (2006) did find cross-modal integration, and in this important respect the studies agree. More generally, the competition/rivalry between explicit and implicit systems in categorization—including the reinforcement gradients and differentials that determine strategy choice—remains an area of scant research that is open to future investigations.

We also considered whether our 8-s timeout for errors changed our paradigm substantively. Eight seconds for reflection might have tilted performance toward dimensional analysis and RB strategies (Maddox, Ashby, Ing, & Pickering, 2004) more than in Maddox et al. (2006), but we did not observe this tilt. In contrast, the timeout might have motivated error avoidance by the undergraduates, fostering adaptive strategies—that is, information integration. We do not exclude this possibility, but we think this effect is small. Using timeouts in other studies, we have replicated canonically many aspects of the RB–II literature—timeouts do not seem to substantially alter the empirical situation.

There are reasons to attribute the performances of the predecisional integrators to processes requiring the basal ganglia, and this attribution has been dominant in the RB–II categorization literature. Many results—from studies using animal lesions, single-cell recording, functional neuroimaging, cognitive testing, and neuropsychological patient groups (Ashby & Ennis, 2006)—link the basal ganglia to II category learning. The striatum (the major input structure in the basal ganglia) receives input from most of cortex, including visual

and auditory association areas (Heimer, 1995). This satisfies one condition for integrating the modalities tested in this study.

However, localizing processes to brain areas is complex. For example, despite evidence that auditory and visual projections to the striatum do have overlap (Arnould, Jeantet, Arsaut, & Demotes-Mainard, 1996; Seimon & Goldman-Rakic, 1985; Yeterian & Pandya, 1998), this may or may not be sufficient to produce the cross-modal effects observed here. Moreover, overlap in area does not imply cross-modal convergence to individual units. Another issue is whether the brain has sufficient cross-modal striatal neurons to produce our integration effects. Probably, there are relatively few multiple-modality neurons (Chudler, Sugiyama, & Dong, 1995). This could explain why only some participants find true integration strategies, and why cross-modal integration is inherently difficult, as is reported here and by Maddox et al. (2006). Another possibility would be for cross-modal convergence at the cortical level, with a multimodal representation projecting to the striatum leading to response selection.

A striatal mechanism for cross-modal integration raises intriguing psychological issues. The cortical projections into the striatum are characterized by massive convergence. Perhaps 30,000 cortical neurons project onto one striatal neuron (e.g., Wilson, 1995). Given this compression of information, the striatum may not remember or store how much activation its neurons received from visual or auditory sources (as predicted, e.g., by Müller's law of specific nerve energies). In fact, the striatum's simple one-layer structure may simply not support dimensional bookkeeping of this kind. The striatum's adaptive function is to produce the final common pathway that leads to the learning and production of adaptive behaviors—little else matters, and possibly the dimensional structure, and even the modality, of the inputs is among the things that do not matter.

Therefore, behavioral signals emerging from the striatum would not be auditory comprehending, or visually focused, or dimensionally attending. These behavioral signals would be nonmodal, even though they were originally produced cross modally. It is interesting that, in the striatal categorization system (which may be the older vertebrate behavioral-categorization system), cognitive evolution may have emphasized nonmodal signals for adaptive behavior within which the dimensional components of the signal are submerged. In other cases, cognitive evolution appears to have placed a premium on maintaining a sharp hold on the dimensional provenance of information, and moreover on strongly analyzing stimuli and responding using one-dimensional rules and dimensional criteria. Certainly this is true of the primate brain, broadly considered (Smith, Beran, et al., 2010; Smith, Crossley, et al., 2012), though perhaps not of all vertebrate species (Smith et al., 2011). However, the present results show clearly that the older, nonmodal system can still be effective in

humans when at need it is called upon, remaining able to integrate information predecisionally, even across different sensory modalities.

Author Note The preparation of this article was supported by Grant No. HD-060563 from NICHD, Grant No. P01 NS044393 from NINDS, and support from the U.S. Army Research Office, through the Institute for Collaborative Biotechnologies under Grant No. W911NF-07-1-0072. We thank our undergraduate research assistants for help with data analysis and data collection.

References

- Arnould, E., Jeantet, Y., Arsaut, J., & Demotes-Mainard, J. (1996). Involvement of the caudal striatum in auditory processing: c-fos response to cortical application of picrotoxin and to auditory stimulation. *Molecular Brain Research*, *41*, 27–35.
- 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*, 442–481. doi:10.1037/0033-295X.105.3.442
- Ashby, F. G., & Ell, S. W. (2001). The neurobiology of human category learning. *Trends in Cognitive Sciences*, *5*, 204–210. doi:10.1016/S1364-6613(00)01624-7
- Ashby, F. G., & Ennis, J. M. (2006). The role of the basal ganglia in category learning. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 46, pp. 1–36). San Diego, CA: Academic Press.
- Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 33–53. doi:10.1037/0278-7393.14.1.33
- Ashby, F. G., & Maddox, W. T. (2010). Human category learning 2.0. *Annals of the New York Academy of Sciences*, *1224*, 147–161.
- Ashby, F. G., Maddox, W. T., & Bohil, C. J. (2002). Observational versus feedback training in rule-based and information-integration category learning. *Memory & Cognition*, *30*, 666–677.
- Ashby, F. G., Queller, S., & Berretty, P. M. (1999). On the dominance of unidimensional rules in unsupervised categorization. *Perception & Psychophysics*, *61*, 1178–1199. doi:10.3758/BF03207622
- Ashby, F. G., & Valentin, V. V. (2005). Multiple systems of perceptual category learning: Theory and cognitive tests. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 547–572). Amsterdam: Elsevier.
- Blair, M., & Homa, D. (2003). As easy to memorize as they are to classify: The 5–4 categories and the category advantage. *Memory & Cognition*, *31*, 1293–1301. doi:10.3758/BF03195812
- Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169–211). Hillsdale, NJ: Erlbaum.
- Chudler, E. H., Sugiyama, K., & Dong, W. K. (1995). Multisensory convergence and integration in the neostriatum and globus pallidus of the rat. *Brain Research*, *674*, 33–45.
- Cook, R. G., & Smith, J. D. (2006). Stages of abstraction and exemplar memorization in pigeon category learning. *Psychological Science*, *17*, 1059–1067.
- DeValois, R. L., & DeValois, K. K. (1988). *Spatial vision*. New York, NY: Oxford University Press.
- Ell, S. W., & Ashby, F. G. (2006). The effects of category overlap on information-integration and rule-based category learning. *Perception & Psychophysics*, *68*, 1013–1026. doi:10.3758/BF03193362
- Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*, 107–140. doi:10.1037/0096-3445.127.2.107
- Feldman, J. (2000). Minimization of Boolean complexity in human concept learning. *Nature*, *407*, 630–633.
- Fukunaga, K. (1972). *Introduction to statistical pattern recognition*. New York, NY: Academic Press.
- Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.
- Heimer, L. (1995). *The human brain and spinal cord* (2nd ed.). New York, NY: Springer.
- Homa, D., Sterling, S., & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Human Learning and Memory*, *7*, 418–439. doi:10.1037/0278-7393.7.6.418
- Jitsumori, M. (1994). Discrimination of artificial polymorphous categories by rhesus monkeys (*Macaca mulatta*). *Quarterly Journal of Experimental Psychology*, *47B*, 371–386.
- Knowlton, B. J., & Squire, L. R. (1993). The learning of categories: Parallel brain systems for item memory and category knowledge. *Science*, *262*, 1747–1749. doi:10.1126/science.8259522
- Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, *53*, 49–70. doi:10.3758/BF03211715
- Maddox, W. T., & Ashby, F. G. (2004). Dissociating explicit and procedural-learning based systems of perceptual category learning. *Behavioural Processes*, *66*, 309–332. doi:10.1016/j.beproc.2004.03.011
- 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*, 650–662. doi:10.1037/0278-7393.29.4.650
- Maddox, W. T., Ashby, F. G., Ing, A. D., & Pickering, A. D. (2004). Disrupting feedback processing interferes with rule-based but not information-integration category learning. *Memory & Cognition*, *32*, 582–591. doi:10.3758/BF03195849
- Maddox, W. T., Bohil, C. J., & Ing, A. D. (2004). Evidence for a procedural-learning-based system in perceptual category learning. *Psychonomic Bulletin & Review*, *11*, 945–952. doi:10.3758/BF03196726
- Maddox, W. T., & Ing, A. D. (2005). Delayed feedback disrupts the procedural-learning system but not the hypothesis-testing system in perceptual category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 100–107. doi:10.1037/0278-7393.31.1.100
- Maddox, W. T., Ing, A. D., & Lauritzen, J. S. (2006). Stimulus modality interacts with category structure in perceptual category learning. *Perception & Psychophysics*, *68*, 1176–1190.
- Minda, J. P., & Smith, J. D. (2001). Prototypes in category learning: The effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 775–799. doi:10.1037/0278-7393.27.3.775
- Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.
- Newell, B. R., Dunn, J. C., & Kalish, M. (2010). The dimensionality of perceptual category learning: A state-trace analysis. *Memory & Cognition*, *38*, 563–581. doi:10.3758/MC.38.5.563
- Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 87–108. doi:10.1037/0278-7393.13.1.87
- Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of multiple-system phenomena in perceptual categorization. *Psychonomic Bulletin & Review*, *7*, 375–402.
- Nosofsky, R. M., Stanton, R. D., & Zaki, S. R. (2005). Procedural interference in perceptual classification: Implicit learning or

- cognitive complexity? *Memory & Cognition*, 33, 1256–1271. doi:10.3758/BF03193227
- Paul, E. J., Boomer, J., Smith, J. D., & Ashby, F. G. (2011). Information-integration category learning and the human uncertainty response. *Memory & Cognition*, 39, 536–554. doi:10.3758/s13421-010-0041-4
- Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, 77, 353–363. doi:10.1037/h0025953
- Reed, S. K. (1978). Category vs. item learning: Implications for categorization models. *Memory & Cognition*, 6, 612–621. doi:10.3758/BF03198251
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605. doi:10.1016/0010-0285(75)90024-9
- Rosseel, Y. (2002). Mixture models of categorization. *Journal of Mathematical Psychology*, 46, 178–210.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461–464. doi:10.1214/aos/1176344136
- Selemon, L. D., & Goldman-Rakic, P. S. (1985). Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. *Journal of Neuroscience*, 5, 776–794.
- Smith, J. D., Ashby, F. G., Berg, M. E., Murphy, M. S., Spiering, B. J., Cook, R. G., & Grace, R. C. (2011). Pigeons' categorization may be exclusively nonanalytic. *Psychonomic Bulletin & Review*, 18, 414–421. doi:10.3758/s13423-010-0047-8
- Smith, J. D., Beran, M. J., Crossley, M. J., Boomer, J., & Ashby, F. G. (2010). Implicit and explicit category learning by macaques (*Macaca mulatta*) and humans (*Homo sapiens*). *Journal of Experimental Psychology: Animal Behavior Processes*, 36, 54–65.
- Smith, J. D., Berg, M. E., Cook, R. G., Boomer, J., Crossley, M. J., Murphy, M. S., & Grace, R. C. (2012). Implicit and explicit categorization: A tale of four species. *Neuroscience & Biobehavioral Reviews*, 36, 2355–2369.
- Smith, J. D., Boomer, J., Zakrzewski, A. C., Roeder, J. L., Church, B. A., & Ashby, F. G. (2014). Deferred feedback sharply dissociates implicit and explicit category learning. *Psychological Science*, 25, 447–457. doi:10.1177/0956797613509112
- Smith, J. D., Chapman, W. P., & Redford, J. S. (2010). Stages of category learning in monkeys (*Macaca mulatta*) and humans (*Homo sapiens*). *Journal of Experimental Psychology: Animal Behavior Processes*, 36, 39–53. doi:10.1037/a0016573
- Smith, J. D., Crossley, M. J., Boomer, J., Church, B. A., Beran, M. J., & Ashby, F. G. (2012). Implicit and explicit category learning by capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, 128, 294–304.
- Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1411–1436. doi:10.1037/0278-7393.24.6.1411
- Smith, J. D., Redford, J. S., & Haas, S. M. (2008). Prototype abstraction by monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: General*, 137, 390–401. doi:10.1037/0096-3445.137.2.390
- Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychonomic Bulletin & Review*, 8, 168–176. doi:10.3758/BF03196154
- Wasserman, E. A., Kiedinger, R. E., & Bhatt, R. S. (1988). Conceptual behavior in pigeons: Categories, subcategories, and pseudocategories. *Journal of Experimental Psychology: Animal Behavior Processes*, 14, 235–246.
- Wilson, C. J. (1995). The contribution of cortical neurons to the firing pattern of striatal spiny neurons. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 29–50). Cambridge, MA: MIT Press.
- Yeterian, E. H., & Pandya, D. N. (1998). Corticostriatal connections of the superior temporal region in rhesus monkeys. *Journal of Comparative Neurology*, 399, 384–402.