

Rich in vitamin C or just a convenient snack? Multiple-category reasoning with cross-classified foods

Brett K. Hayes · Hendy Kurniawan · Ben R. Newell

Published online: 3 November 2010
© The Psychonomic Society 2010

Abstract Two studies examined multiple category reasoning in property induction with cross-classified foods. Pilot tests identified foods that were more typical of a taxonomic category (e.g., “fruit”; termed ‘taxonomic primary’) or a script based category (e.g., “snack foods”; termed ‘script primary’). They also confirmed that taxonomic categories were perceived as more coherent than script categories. In Experiment 1 participants completed an induction task in which information from multiple categories could be searched and combined to generate a property prediction about a target food. Multiple categories were more often consulted and used in prediction for script primary than for taxonomic primary foods. Experiment 2 replicated this finding across a range of property types but found that multiple category reasoning was reduced in the presence of a concurrent cognitive load. Property type affected which categories were consulted first and how information from multiple categories was weighted. The results show that multiple categories are more likely to be used for property predictions about cross-classified objects when an object is primarily associated with a category that has low coherence.

Keywords Induction · Categorization · Cross-classification

Introduction

Most things belong to more than one category. *Hilary Clinton* is Secretary of State, the wife of a former President, a Democrat, an American and a woman. An *apple* is a fruit

but can also be thought of as a snack or lunch food. Such cross-categorization raises an interesting question about property inference. To what extent do we consider multiple category membership when making inferences about cross-classified items? This question is important because whether we consider multiple alternative categories or just one category determines the nature of our inference. If asked to predict Hilary Clinton’s views on the use of military force against Iran, for example, we might arrive at different predictions depending on whether we consider her primarily as Secretary of State, as a Democrat or as a woman.

Previous research has produced mixed results regarding the extent to which people consider multiple categories when making inferences about cross-classified stimuli. Early work with members of cross-classified social categories (e.g., a *Chinese woman*) has suggested that activation of one category alternative inhibits consideration of the other alternative (Macrae, Bodenhausen, & Milne, 1995). Hence, inductive inferences about cross-classified items will be based on the single category that seems most relevant to the prediction at hand (Nelson & Miller, 1995). In this respect, the way people reason about cross-classified items has been seen to parallel reasoning about items whose category membership is ambiguous. For example, a physician who sees a patient with a skin blemish may believe that it is most likely a harmless sun-spot, but may also recognise there is some chance it belongs to an alternative diagnostic category (e.g., skin cancer). A considerable body of evidence shows that in cases of ambiguous categorization, people consider only the most likely category when making inductive predictions (Murphy & Ross, 2007, but see Hayes & Newell, 2009).

Murphy and Ross (1999), on the other hand, found some evidence for multiple-category inference with cross-classified items, but only when the experimental instructions and procedure encouraged such reasoning. Murphy

B. K. Hayes (✉) · H. Kurniawan · B. R. Newell
School of Psychology, University of New South Wales,
Sydney, NSW 2052, Australia
e-mail: B.Hayes@unsw.edu.au

and Ross used an information search task where people could consult multiple categories when making inferences about a novel property of cross-classified foods. For example, when asked to predict how often an apple was “low in essential amino acids” participants had the option of examining property prevalence in at least two relevant categories (“fruit” and “snacks”). When the experimenter provided a practice trial demonstrating that more than one category could yield multiple pieces of relevant property information, people frequently consulted multiple categories. Without this priming, however, people usually made predictions based on a single category. Murphy and Ross concluded that multiple category reasoning was likely to occur only when attention was drawn to the relevance of the various category alternatives to the prediction at hand.

Category coherence and induction with cross-classified items

Our studies are concerned with what is likely to be an even more important determinant of multiple category reasoning: the perceived coherence of the various categories to which an item belongs. Category coherence refers to the extent to which category features and/or exemplars are seen to go together as a function of background knowledge about their causal origin or purpose (Haslam, Rothschild, & Ernst, 2000; Ross & Murphy, 1999). Coherence generally promotes inductive generalization. When people are aware of the causal basis for category membership they are more likely to generalize category properties to a novel instance than when category membership is defined purely on the basis of similarity (Rehder, 2009). People are also more likely to make property inferences based on categories with high as opposed to low levels of coherence (Patalano & Ross, 2007).

Patalano, Chin-Parker and Ross (2006) found that the relative coherence of category alternatives influenced inferences about instances belonging to two or more social categories. For example, people were told that feminists (high coherence category) preferred Coke to Pepsi but waiters (low coherence category) showed the opposite pattern. When asked to predict the drink preferences of “feminist supporters who were waiters”, most people made predictions in line with the more coherent component.

Patalano et al. (2006) also examined multiple-category reasoning using an information search procedure where participants were asked to make a prediction about an individual (e.g., likelihood of them preferring Coke to Pepsi) after examining information from up to four social categories to which the individual belonged. Overall, participants showed some evidence of multiple category reasoning, typically examining two out of four relevant

categories before making a prediction. Critically, participants examined categories rated as having high coherence before they examined low coherence categories.

These data highlight that the relative coherence of category alternatives is important when people make inferences about cross-classified items. Our studies aimed to extend the study of the role of coherence in multiple category reasoning in two ways. First, we addressed an important limitation of the information search study by Patalano et al. In that study, participants were told that a given item *was* a member of all of the visible categories. This explicit instruction may have primed participants to search multiple relevant categories (cf. Murphy & Ross, 1999, Experiment 4), leading to an inflated estimate of multiple-category reasoning. Our studies, therefore, used an information search paradigm in which participants had to rely on their background knowledge of the target objects (cross-classified foods) to decide which categories should be examined before making an inductive prediction. We believe that this scenario is closer to situations likely to be found outside of the laboratory, and hence should yield more generalizable findings regarding patterns of single and multiple category reasoning.

Second, and more importantly, we examine a context where multiple-category reasoning seems more likely to occur, namely, when a cross-classified item is most strongly associated with a category that has relatively low coherence. In everyday reasoning cross-classified items may often be more strongly associated with one kind of category than another. Ross and Murphy (1999), for example, have shown that people recognise that many foods belong to both folk taxonomic (e.g., fruit, meat, dairy) and script based categories organised around the time or situation in which a food is eaten (e.g., snacks, breakfast foods, dinner foods). However, it seems likely that individual foods will be more strongly associated with one type of category than the other. In support of this notion, Murphy and Ross (1999) found that certain cross-classified foods were preferentially associated with one type of category. For example, an *apple* was rated as a more typical example of the category “fruit” than of the category “snack foods”, even though it was viewed as belonging to both categories. Conversely, cornflakes are likely to have a stronger association with the script category of “breakfast food” than with the taxonomic category of “breads/grains”, but still be seen as belonging to both. Following Murphy and Ross (1999), we refer to these different types of cross-classified instances as “taxonomic primary” and “script primary”.

This distinction is likely to be critical in determining whether people consult multiple categories when reasoning about cross-classified items. If a primary category for a target instance is highly coherent then it seems likely that

property inferences about the target may be based on membership of this category alone, without considering cross-classified alternatives. This is suggested by Patalano et al. (2006) who found that once information from a highly coherent category was accessed, other category alternatives were less likely to be examined. Alternately, if the primary category for a cross-classified item is lacking in coherence then it seems likely that an individual *will consult and use information from other relevant categories* when making a prediction.

The main aim of our experiments was to test these predictions about the effects of primary and secondary category coherence on multiple category reasoning in the food domain. We first identify two different kinds of cross-classified foods: those seen as better examples of a taxonomic category than a script category (taxonomic primary or “tax primary” foods) and those seen as better examples of script than taxonomic categories (script primary foods). Second, we test the prediction that people generally perceive taxonomic food categories to be more coherent than script categories. This seems likely given previous work (e.g., Barr & Caplan, 1987; Gentner & Kurtz, 2005) which has found that categories based on extrinsic relations (e.g., “items associated with dining out at a restaurant”) are seen as less coherent than taxonomic categories.

The findings of these pretests leads to the key prediction tested in Experiments 1 and 2 that people will be more likely to engage in multiple category inference for script primary foods than for taxonomic primary foods. Experiment 2 examined additional factors that may moderate the effects of primary and secondary category coherence, including the availability of processing resources, and the nature of the properties being predicted.

Experiment 1

This study proceeded in two stages. The first stage involved extensive pilot testing of the food items to (a) identify cross-classified food exemplars that differ in the extent to which they are seen as associated with taxonomic or script categories, and (b) confirm that taxonomic food categories are seen as more coherent.

The second stage was the main experiment which employed an information search and prediction paradigm similar to Murphy and Ross (1999, Experiment 5). On each trial participants were given a food exemplar (e.g., toast) and asked to make a prediction about a novel quantitative property (e.g., “What percentage of the time does the food induce gluconeogenesis?”). They could learn about the distribution of this property within various food categories by clicking on up to four category labels. On target trials two

of these categories were the relevant taxonomic (e.g., breads/grains) and script (e.g., breakfast) alternatives, and two were irrelevant (e.g., for toast, “meat” and “dairy food” were irrelevant distracters). Each relevant category contained different numerical property information. Hence, if someone searched both relevant categories *and* integrated the relevant property information they should make a prediction that was between the two estimates from the relevant categories. We predicted that people would be more likely to search multiple relevant categories and be more likely to integrate property information across these categories for script primary than for taxonomic primary foods.

Method

Participants

The participants were 39 first-year psychology students who participated for course credit. One participant was eliminated because she appeared to ignore the instructions about only consulting relevant categories and clicked all four categories on every trial. An additional 32 undergraduates made pilot typicality ratings of foods and a further 20 undergraduates rated the coherence of taxonomic and script categories. No one participated in both the pilot and the main studies.

Typicality ratings

The goal of this pilot test was to identify cross-classified foods that were seen as better examples of taxonomic than script categories (tax primary) or vice versa (script primary). The use of typicality ratings as a measure of the relative association of an item with alternate categories is somewhat novel, but was motivated by previous work which shows that typical items are more often produced in response to category names than atypical items (Mervis, Catlin, & Rosch, 1976). In this respect, an item’s typicality reflects its relative availability within a category.

Typicality ratings were obtained for 30 different foods. Approximately half were adapted from Ross and Murphy (1999, Experiment 2) and others were added to suit local dietary patterns. Pilot participants rated the typicality of each food on a seven-point scale (1 = “not at all typical”, 7 = “highly typical”) for five folk-taxonomic categories (“fruit”, “meats”, “breads/grains”, “dairy”, and “vegetables”) and five script based categories (“breakfast”, “lunch”, “dinner”, “snacks” and “dessert”).

To be viewed as cross-classified an item had to have a mean typicality rating of at least 4.0 in at least one taxonomic category and one script category. Items were identified as taxonomic or script primary if they satisfied this condition *and* had a typicality rating above 5.5 for one

of the category alternatives but not the other. For example, “apples” received a typicality rating of 7.0 for the taxonomic category “fruit”, and a rating of 5.25 for the script category “snacks”, and therefore was designated as a taxonomic primary food. Twenty items met these criteria (see Table 1). Typicality ratings for the primary category were reliably higher than those for the secondary category for both taxonomic primary items, $t(9) = 7.18, p < 0.001$, and script primary items, $t(9) = 4.3, p < 0.01$. We also examined whether the difference in typicality ratings between primary and secondary categories varied with category type. The difference in typicality ratings between taxonomic primary and script secondary categories was compared to the corresponding difference between ratings for script primary and taxonomic secondary categories. The size of this difference did not vary with category type, $t(18) = 1.4, p = 0.19$. Participants were just as good at discriminating between taxonomic primary and script secondary categories for foods as they were at discriminating between script primary and taxonomic secondary categories.

Category coherence ratings

Twenty participants rated the five taxonomic and five script categories on scales of within-category similarity and infor-

mativeity. For similarity, participants rated each category on a nine-point scale regarding “how similar or diverse are the members of each food category?” (1 = diverse/differing, 9 = uniform/similar). For informativeness, participants rated each category on a nine-point scale of whether “knowing that something belongs in the category tells us a lot about that food” (1 = uninformative; 9 = informative). We also obtained category familiarity ratings (e.g., “On average, within a single week how often do you think Australians would eat something from this category/group of foods?”; 1 = not at all, 9 = very frequently) to check whether coherence was confounded with personal experience of food categories.

Table 2 shows the mean ratings for each food category. As expected, taxonomic categories received higher ratings of within-category similarity ($M = 5.65$) and informativeness ($M = 6.72$) than script categories ($M_s = 3.08; M = 4.24$, respectively), $p_s < 0.001$. The two category types were rated as equally familiar (taxonomic: $M = 6.44$; script; $M = 6.90$), $t(19) = 1.88, p = 0.08$. Item-wise correlations between ratings were calculated across the ten categories. In line with previous findings (e.g., Haslam et al., 2000; Patalano et al., 2006), ratings of similarity and informativeness were highly correlated, $r(9) = 0.92, p < 0.001$, but familiarity was not correlated with similarity, $r(9) = -0.28, p = 0.44$, or informativeness, $r(9) = -0.49, p = 0.1$.

Table 1 Taxonomic primary and script primary cross-classified foods with corresponding categories and typicality ratings

Item	Taxonomic category	Taxonomic typicality	Script category	Script typicality	Item coherence (within primary food type)
Taxonomic primary items					
Apples	Fruit	7.0	Snacks	5.25	Low
Bacon	Meat	6.56	Breakfast	5.50	Low
Baked carrots	Vegetables	6.94	Dinner	5.25	High
Celery Sticks	Vegetables	6.38	Snacks	4.88	High
Cheese	Dairy foods	6.44	Snacks	4.00	High
Corn	Vegetables	6.69	Dinner	4.38	High
Grapes	Fruit	6.81	Snacks	4.81	Low
Lamb	Meat	6.69	Dinner	5.25	Low
Meatballs	Meat	6.69	Dinner	5.50	Low
Salami	Meat	6.5	Dinner	5.50	Low
Script primary items					
Cornflakes	Breads/grains	4.63	Breakfast	6.88	High
Toast	Breads/grains	5.06	Breakfast	6.75	High
English Muffin	Breads/grains	4.80	Breakfast	6.20	High
Custard	Dairy	5.00	Dessert	5.56	High
Chicken Wrap	Meat	5.40	Lunch	6.30	Low
Ice-cream	Dairy	5.41	Dessert	6.88	High
Kebab	Meat	5.50	Lunch	6.00	Low
Pide	Breads/grains	5.00	Lunch	5.70	Low
Potato	Vegetables	5.50	Dinner	5.70	Low
Sandwich	Breads/grains	4.56	Lunch	6.63	Low

Table 2 Mean pretest ratings for food categories

Category	Category type	Informativeness	Within-category similarity	Familiarity
Dairy	Taxonomic	6.85	5.5	6.7
Vegetable	Taxonomic	7.05	5.4	4.85
Fruit	Taxonomic	6.65	5.05	5.55
Meat	Taxonomic	6.3	5.65	7.85
Bread/Grains	Taxonomic	6.75	6.65	7.25
Snacks	Script	4.85	3	7.5
Breakfast	Script	5.5	4.25	5.15
Dinner	Script	3.9	2.35	7.35
Desserts	Script	5.4	3.25	6.8
Lunch	Script	3.55	2.55	7.7

Materials and procedure for information search

The main experiment consisted of 26 induction trials presented on a 19" computer screen controlled by Revolution v2.0 software. Twenty trials were targets in which a specific food (e.g., toast) was presented and the participant was asked to make a prediction about a novel numerical food property (see Fig. 1 for a trial example). Properties were generated from Wikipedia files about food or were invented by the authors and related to the food's biochemical composition or chemical reactions to its consumption (e.g., "What percentage of the time does this food contain beta-globulin?"; "What percentage of the time does the food induce gluconeogenesis?"). All property predictions required a response on a 0–100 scale.

In order to make a prediction, participants were instructed to consult up to four categories presented on the lower half of the screen. On all target trials two categories were relevant to the given food exemplar and two were irrelevant. For half of the target trials the given food was one of the taxonomic primary foods (e.g., "apple") identified in pilot testing while for the remaining trials it was a script primary food (e.g., "toast"). Clicking with the mouse on a category label revealed two statements about numerical category properties; one of these was always the critical property for which a prediction was required and the other concerned an irrelevant property (see Fig. 1). When participants clicked on the primary category they saw a probability rating between 45% and 55% for the critical property. If they clicked on the secondary category then they were given information about the critical property that differed by 40% from the corresponding value associated with the primary category. On half the trials this value was above that for the primary category ("high trials") while for the remainder it was below that value ("low trials"). Integration of property information across categories was indicated by a numerical response (entered via the keyboard)

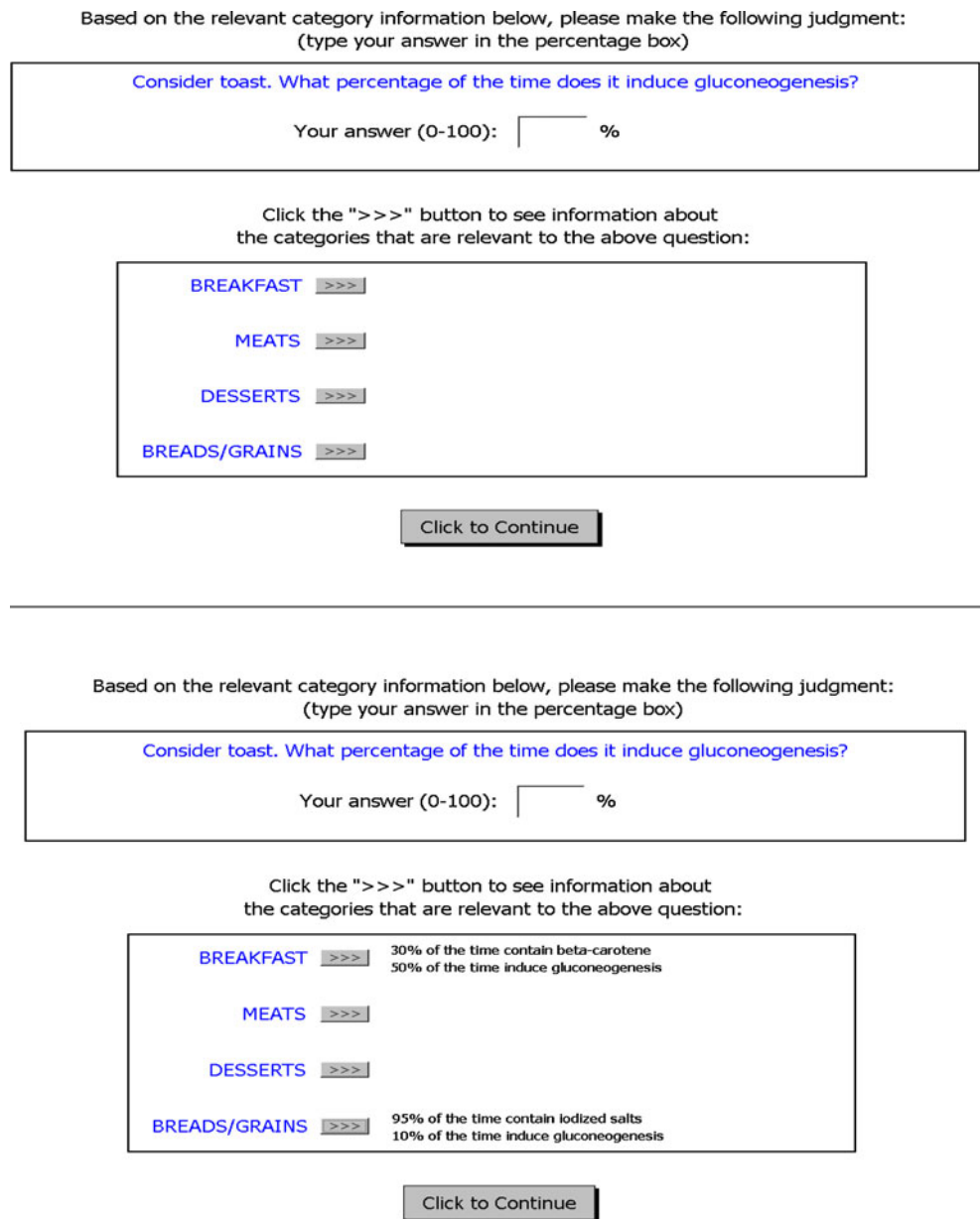
that fell between values associated with the primary and secondary categories (referred to as an "in-between" response). For example, in Fig. 1 the relevant property figure from the primary category of "breakfast foods" was 50%, while the relevant figure from the secondary category "breads/grains" was 10%. To be counted as an in-between response the property estimate had to be at least five percentage points away from these endpoints and in the predicted direction (i.e. had to fall in the range of 15–45%). A similar scoring rule was used by Murphy and Ross (1999).

Clicking on irrelevant categories also revealed two property statements, one of which concerned the critical property. For these categories, however, the value for the critical property never fell between the values associated with the two relevant categories. Hence an in-between response could only result from consulting *relevant* categories. On the six filler trials we presented foods that, in pilot testing, were rated as typical of only one of the category types (i.e. were not cross-classified). Data for these trials were not analysed.

All participants were tested individually. Participants were told to consult "all categories that were relevant" before making a prediction but that examining irrelevant categories would not be helpful. Murphy and Ross (1999) found that the information search procedure was sensitive to demand characteristics such as pre-test demonstrations of the results of clicking on each category. Hence, we gave written descriptions but did not demonstrate the kinds of property information that would be revealed as each category was clicked.

Trials were self-paced with a 2 s blank screen inter-trial interval. On a given trial the order in which categories appeared on the screen was randomized with the constraint that the two relevant categories never appeared in adjacent positions. Within each relevant category the order of critical and irrelevant property information was randomized. The presentation order of taxonomic primary, script primary and filler items was also randomized.

Fig. 1 Example of an induction trial display before (*upper panel*) and after (*lower panel*) a participant has searched relevant categories (Note: In the *lower panel* the relevant property figures are “50%” from breakfast foods and “10%” from breads/grains. An “in-between response” for this item would be a property estimate between 15% and 45%)



Results

The main outcome measures were the frequency with which multiple categories were examined and the percentage of trials where the numerical feature prediction was in-between those associated with the two relevant categories.

Category examination

The distribution of category examination is given in Fig. 2. Overall, only one category was consulted on most trials ($M_{CLICKS\ PER\ TRIAL} = 1.53$), a result that closely parallels the most similar study of Murphy and Ross (1999, Experiment 5). Participants rarely examined three or four

categories (less than 8% of trials in total). When only a single category was consulted, it was most often taxonomic (79% of single-category trials), but the likelihood of consulting a single taxonomic category differed across food type, (tax primary = 93% of single category trials vs. script primary = 67% of single category trials), $F(1, 33) = 28.4$, $p < 0.001$, partial $\eta^2 = 0.46$.

As predicted, however, people were more likely to examine two or more categories for script primary (47.4% of all target script trials) than taxonomic primary foods (29.4% of all target taxonomic trials), $F(1, 37) = 20.4$, $p < 0.001$, partial $\eta^2 = 0.36$. Note that for both taxonomic and script primary foods, the categories examined on the first or second selection were almost always those “relevant” to the food exemplar (94% of occasions). This means that the

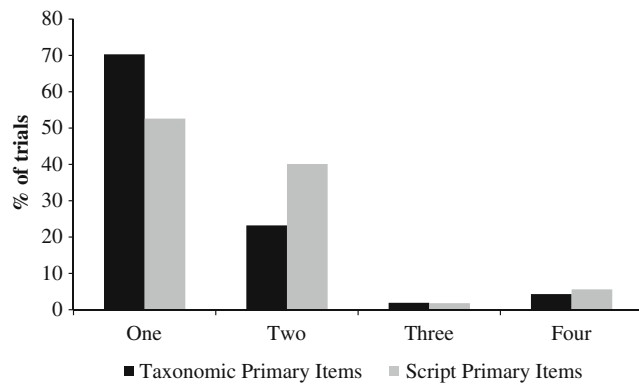


Fig. 2 Percentage of target induction trials in which one, two, three or four categories were examined in Experiment 1

pattern of category examination in Fig. 2 reflects the extent to which people searched for property information from multiple relevant categories. For taxonomic primary items people most often only examined the relevant taxonomic category and then ended their search. For script primary items they were more likely to consult both taxonomic and script categories.

An examination of individual profiles reinforces this picture. Participants were classified as consistently consulting multiple categories if, for a given food type, they clicked on more than one category on more than 50% of target trials. Of the 38 participants, nine consistently consulted multiple categories for both tax primary and script primary items. Notably, an additional nine participants consistently consulted a single category for tax primary foods but consulted multiple categories for script primary foods. No participants, however, showed the reverse pattern.

Although we did not make strong predictions about the order in which categories would be examined, our account implies that people should generally examine the primary category for a given item before examining a secondary category. This prediction was only partially correct. As expected, order of examination was affected by item type, with the proportion of trials where the script category was the first one examined higher for script primary ($M = 41\%$) than for tax primary foods ($M = 15\%$), $F(1, 36) = 21.56$, $p < 0.001$, partial $\eta^2 = 0.31$. Note, however, that even for script primary foods, taxonomic categories were more often examined before script categories. There appeared to be a general bias towards early examination of taxonomic categories for all food types.

To further examine the relationship between order of examination and number of categories examined, we computed the conditional probabilities that a participant would examine other categories after they had examined either a taxonomic or a script-based category (i.e. the proportion of trials in which either a taxonomic or script category was clicked first and other categories subsequently

examined was divided by the total number of trials where a taxonomic or script category was clicked first). For tax primary items, the probability that participants would continue to search other categories after examining a script category (0.54) was significantly higher than the probability of continued search after examining a taxonomic category (0.27), $t(35) = 4.17$, $p < 0.001$. For script primary items, the corresponding difference in conditional probabilities (0.46 after examining a script category, 0.37 after examining a taxonomic category) approached significance, $t(37) = 1.82$, $p = 0.08$

In-between responding and inductive prediction

An in-between response involved cases where two or more categories were examined *and* a property estimate was given that lay between the end-points associated with the two relevant categories. A preliminary analysis found no differences in the proportion of in-between responses for trials with “high” and “low” estimates ($F < 1.0$) so responses on these trials were collapsed. In-between responses were only made on a subset of trials where more than one category was examined (24% of all target trials). As predicted, however, more in-between responses were made for script primary (27% of target trials) than tax primary items (21% of target trials), $F(1, 37) = 4.81$, $p = 0.03$, partial $\eta^2 = 0.12$.

A minority of individuals consistently made in-between predictions (i.e. on more than half the trials where multiple categories were clicked) for both food types ($n = 8$; 21%). Some individuals consistently made multiple category predictions for script but not tax primary items ($n = 3$; 8%), but no one showed the opposite pattern.

The relative weighting of property information from taxonomic and script categories was examined in an analysis of the numerical predictions given for in-between responses. If participants gave equal weight to the two category types then their predictions should be close to the mean of the property values associated with them. For each in-between response we calculated the difference between this mean value and the numerical prediction given. Difference scores for each food item were then averaged across participants. We reversed the sign of difference scores for items where the property value associated with the script category was higher than that of the taxonomic category, so that all difference scores were scaled in the same direction. As a result, a positive difference score reflected a bias towards values associated with the taxonomic category, whereas a negative difference reflected a bias towards the script category. Equal weighting of information from each category type was indicated by a difference score of zero. The mean difference score for predictions based on tax primary and script primary items was computed and compared against a value of zero in single-sample t -tests. For taxonomic-primary items, in-between predictions were biased

in the direction of the value associated with the taxonomic category ($M_{\text{DIFFERENCE SCORE}} = 3.2$, $t(9) = 5.02$, $p < 0.001$). For script-primary items there was no such bias, suggesting that people gave roughly equal weight to evidence from script and taxonomic categories, ($M_{\text{DIFFERENCE SCORE}} = 0.24$, $t(9) = 0.28$, $p = 0.79$).

Coherence effects within category types

Our pretesting established clear differences in the perceived coherence of script and taxonomic categories. For example, Table 2 shows that “breakfast foods”, the category that was rated as *most* coherent on informativeness and within-category similarity scales amongst script categories, received lower similarity and informativeness ratings than “fruit”, which was judged as the *least* coherent of the taxonomic categories. Nevertheless, this does not rule out the possibility that other kinds of differences between script primary and tax primary items may have contributed to differences in multiple category reasoning.

To better understand the extent to which reasoning was affected by coherence, over and above differences in category type (script vs. taxonomic), we examined the relative coherence of food categories *within* each of the taxonomic and script groupings. Categories were ranked according to their mean pretest ratings of informativeness and within-category similarity to identify the two taxonomic categories (“dairy”, “vegetables”) and two script categories (“breakfast foods”, “deserts”) that had the highest coherence ratings within their category types and the corresponding categories that received the lowest ratings (taxonomic: “fruit”, “meat”; script: “dinner foods”, “lunch foods”).¹ Each of the 20 food items in Table 1 was then classified as having either high or low coherence depending on the relative coherence of its primary category (with respect to other categories of the same food type). The right-hand column in Table 1 shows the resulting coherence classification of individual items. The mean proportion of trials where two or more categories were examined and proportion of trials where an in-between prediction was made were then compared for high and low coherence items within each food type. Consistent with our predictions about coherence effects, people were more likely to examine multiple categories ($t(37) = 4.59$, $p < 0.001$) and make an in-between response ($t(37) = 3.30$, $p = 0.002$) when a script primary item had relatively *low* coherence than when it had high coherence. In other words, category

coherence affected multiple category reasoning *within* the set of script primary items. No differences on multiple category reasoning measures were found for low and high coherence taxonomic items (p 's > 0.45).

Discussion

Like Murphy and Ross (1999), Experiment 5 we found a modest level of multiple category reasoning overall. The critical novel finding, however, was that multiple category reasoning depended on the relative coherence of categories associated with a cross-classified instance. People were more likely to consult information from multiple categories about cross-classified foods that were primarily associated with script categories than foods primarily associated with taxonomic categories. This same pattern was found in the likelihood that people would *use* multiple category information in deriving a prediction. These data support the idea that if people access a highly coherent category early in the induction process then they are likely to terminate their search for further property information and make a prediction based on this category. If, however, people access a more loosely organized category they are more likely to use multiple category reasoning.

Although we found that patterns of multiple-category reasoning depended on whether a food was primarily associated with a taxonomic or script category, there were some results which also suggested a general bias in favour of predictions based on taxonomy. People consulted taxonomic categories before script categories, even on script primary items. Moreover, for taxonomic primary items where both taxonomic and script alternatives were consulted, people made numerical predictions that were biased towards the information associated with the taxonomic alternative. This could reflect a general preference for taxonomic information when making predictions about foods. Alternately, it could be an artifact caused by the use of intrinsic properties like biochemical composition, which invite taxonomic generalization (cf. Ross & Murphy, 1999). These alternative explanations are examined in Experiment 2.

Experiment 2

Experiment 1 showed that people were more likely to consult multiple categories when making predictions about script primary foods. A limitation of that study, however, was that predictions always involved biochemical properties. Previous work has shown that different types of properties can selectively activate different kinds of category knowledge (Heit & Rubenstein, 1994; Nguyen &

¹ Table 2 shows that “bread/grains” received the highest coherence score (mean of informativeness and similarity ratings) among taxonomic categories. However, because no food item was primarily associated with breads/grain, this category was not used in the analysis of coherence effects within food type.

Murphy, 2003). In folkbiology in general (e.g., Shafto et al., 2007), and foods in particular (Ross & Murphy, 1999), intrinsic object properties like anatomical structure are typically generalized along taxonomic lines. Hence the use of biochemical properties may have biased the findings in favor of taxonomic categories and underestimated the extent to which people activate script categories in multiple category reasoning.

Experiment 2 addressed this issue by asking different groups to make predictions about either structural/biochemical properties, situational properties involving the details of where or when a food is eaten (e.g., “percentage of the time it is served at the Caribana Festival in Cuba”) or blank properties (e.g., “has property Q2”). In our studies, script-based categories were created by grouping together foods that were eaten at a particular time or location. It seems reasonable to expect more examination of script categories for induction where situational properties are the focus. In the case of blank properties people may be unsure whether the property in question is more closely related to taxonomic or script based alternatives so we may see relatively high levels of multiple category reasoning in this condition. Alternately, the blank condition may reveal a default bias in preference for taxonomic or script-based induction.

Experiment 2 examined a second important issue regarding use of multiple categories in property prediction, namely, the availability of processing resources. It has been suggested that reasoning with multiple categories involves greater demands on working memory than reasoning from a single certain category (e.g., Ross & Murphy, 1996). To date, however, this hypothesis has not been tested. Therefore, we presented the induction task from Experiment 1 under conditions of low and high concurrent memory load. Adding a concurrent load should reduce the likelihood that people will search more than one category for property information. Given that more multiple category reasoning was found for script primary items in Experiment 1, it was expected that the memory load manipulation would have a greater impact on these items.

Method

Participants included 128 undergraduate psychology students who took part for course credit. One participant clicked all four categories on every trial and was eliminated from analyses.

Design and procedure

The experiment had a 3 (property) \times 2 (memory load) \times (2) (item type: tax primary vs. script primary) design with repeated measures on the last factor. The general procedure was the same as in Experiment 1 but with additional

between-group manipulations. Those in the biochemical property condition made predictions about the same properties as in Experiment 1. Those in the situational condition made predictions about properties that related to the place or occasion of food consumption (see Appendix). Those in the blank property condition made predictions about properties denoted by random letter-number combinations.

Before the start of each induction trial participants were presented with the first part of a dot-memory task adapted from De Neys (2006). Participants saw a 12 cm² square in the center of the screen divided into a 3 \times 3 grid. Those in the low memory load condition saw three black dots located in three cells of the grid for 3 s. In this condition, the dots were always presented in a vertical or horizontal line. De Neys (2006) showed that this condition places minimal demands on memory. In the high load condition five dots were presented in random locations across the grid. All participants were told to remember the location of the dots while they completed the subsequent induction problem. After each induction trial a blank 3 \times 3 grid was presented and participants had to click the mouse to indicate the position of each of the dots presented earlier. Dot recall was self-paced. A dot memory trial was scored as correct only if all dots were recalled in the correct location.

Results

Performance on the dot-memory task was good in both low and high load conditions ($M = 92\%$ items correct, $SD = 8.1\%$), with every participant completing at least 70% of items correctly. All participants were therefore deemed to have devoted sufficient attention to this secondary task, and data for all trials were included in subsequent analyses.

Category examination

Induction performance was analyzed in the same way as in Experiment 1. Again, overall, only one category was usually consulted before a prediction was made ($M_{\text{CLICKS PER TRIAL}} = 1.54$). When only a single category was consulted, it was most often taxonomic (75% of single-category trials), but the likelihood of consulting a single taxonomic category differed significantly across food type, (tax primary = 91% of single category trials vs. script primary = 59% of single category trials), $F(1, 106) = 140.07$, $p < 0.001$, partial $\eta^2 = 0.57$.

Replicating the main finding of Experiment 1, however, consultation of multiple categories depended on item type. Multiple categories were consulted more often for script primary ($M = 39\%$ of total trials) than tax primary foods ($M = 27\%$ of total trials), $F(1, 121) = 60.32$, $p < 0.001$, partial $\eta^2 = 0.33$. As predicted, this effect interacted with

memory load, $F(1, 121) = 4.35, p = 0.04, \text{partial } \eta^2 = 0.04$. Figure 3 shows that, under high load, consultation of multiple categories for script primary foods was reduced, but there was little change to information search for tax primary items. When blank properties were used people tended to examine multiple categories more often than in the structural or situational property conditions, but this effect was not reliable, $F(2, 121) = 2.57, p = 0.08$.

Again it was instructive to examine individual patterns of information search. Thirty-eight participants (30% of the total sample) consistently examined multiple categories for both types of items. Once again there were more participants who consulted multiple categories for script items and single categories for taxonomic items ($n = 17; 13\%$) than who showed the opposite pattern ($n = 2; 1.6\%$).

When we examined the order in which categories were examined, there was a bias towards the early examination of taxonomic categories, which were examined before script categories on 65% of trials overall. As in Experiment 1, however, order of examination depended on food type, with script categories more likely to be examined first for script primary items ($M = 47\%$) than for tax primary items ($M = 16\%$), $F(1, 121) = 229.67, p < 0.001, \text{partial } \eta^2 =$

0.65. A novel finding was that order of examination was also influenced by property type, $F(2, 121) = 6.26, p = 0.003, \text{partial } \eta^2 = 0.09$. Planned contrasts confirmed that script categories were more likely to be consulted before taxonomic categories when predictions were required about situational properties ($M = 38\%$) than in the blank ($M = 30\%$) or biochemical conditions ($M = 27\%$), ($p = 0.001$). The interaction between food type and property type was marginally significant, $F(2, 121) = 3.04, p = 0.05, \text{partial } \eta^2 = 0.05$. The difference between early examination of script categories for tax and script primary items tended to be larger when situational properties were used (script primary $M = 0.56$; tax primary $M = 0.19$), as compared with the biochemical and blank conditions (script primary $M = 0.43$; tax primary $M = 0.15$). Memory load had no significant effects on order of examination (F 's < 2.0).

Again, to examine the relationship between order of examination and number of categories examined, we compared the conditional probabilities that a participant would examine other categories after they had examined either a taxonomic or a script-based category. For tax primary items, the probability that participants would continue to search other categories after examining a script category (0.49) was significantly higher than the probability of continued search after examining a taxonomic category (0.29), $t(102) = 5.28, p < 0.001$. The corresponding probabilities for script primary items showed a similar pattern (0.44 vs. 0.39), but this difference was not significant, $t(122) = 0.97, p = 0.17$.

In-between responding and inductive prediction

Figure 4 shows the percentage of trials where people consulted more than one category and made an in-between prediction. The percentage of such trials was slightly lower overall than in Experiment 1 (by 6%), most likely due to the imposition of the secondary memory task. Notably though the level of multiple category reasoning still differed according to food type. More in-between predictions were made for script primary ($M = 21\%$ of total script trials) than for tax primary foods ($M = 15\%$ of total tax trials), $F(1, 121) = 14.71, p < 0.001, \text{partial } \eta^2 = 0.11$. There were no main effects or interactions involving the load or property manipulations (F 's < 1.6).

Patterns of individual data for consistent in-between predictions generally followed the trends found for information search. A minority of participants consistently made in-between predictions for both food types ($n = 7; 12\%$). More participants consistently made multiple category predictions for script primary items and single category

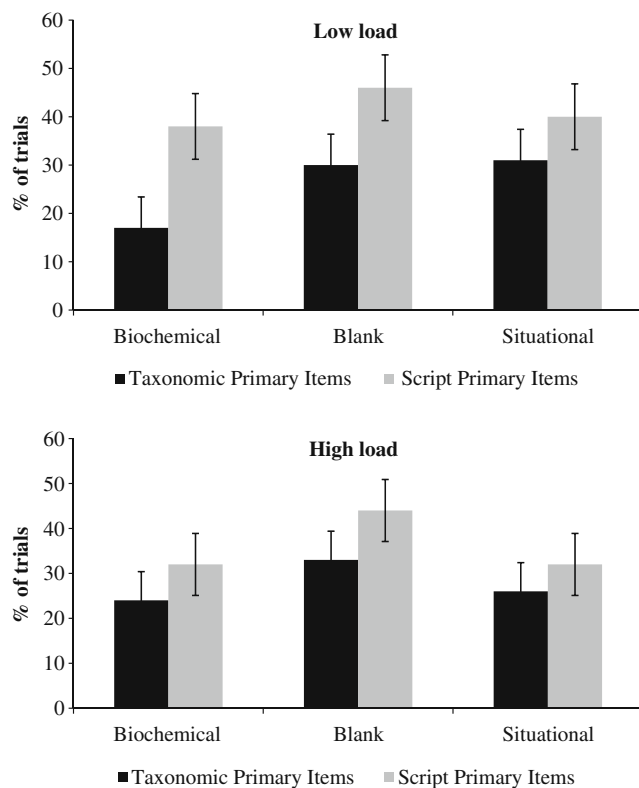


Fig. 3 Percentage of target trials where multiple-categories (2, 3, or 4) were examined in Experiment 2

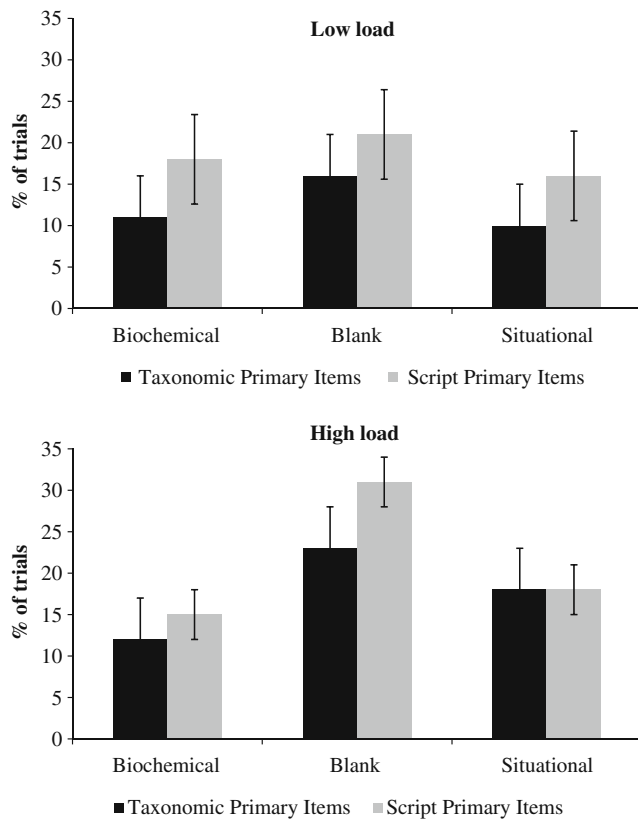


Fig. 4 Percentage of target trials where multiple-categories were examined and an in-between prediction was made in Experiment 2

predictions for taxonomic items ($n = 7$; 6%) than the other way around ($n = 2$; 1.6%).

We examined the relative weighting of property information in numerical predictions from taxonomic and script category alternatives, employing the same approach as in Experiment 1 but examining property and load conditions separately. When biochemical properties were used, we replicated the Experiment 1 findings. Numerical predictions were biased in the direction of the taxonomic category for taxonomic primary items ($M_{\text{DIFFERENCE SCORE}} = 5.73$, $t(9) = 9.42$, $p < 0.001$), but were unbiased (i.e. equal weighting given to taxonomic and script information) for script primary items ($M_{\text{DIFFERENCE SCORE}} = 1.46$, $t(9) = 1.24$, $p = 0.25$). A similar pattern was found when blank properties were used (tax primary items: $M_{\text{DIFFERENCE SCORE}} = 2.89$, $t(9) = 2.78$, $p = 0.02$; script primary items: $M_{\text{DIFFERENCE SCORE}} = -0.25$, $t(9) = -0.3$, $p = 0.77$). Notably, the bias favouring information from taxonomic categories for tax primary items disappeared when situational properties were used. In this group, information from each category type was given equal weighting in all numerical predictions (tax primary items: $M_{\text{DIFFERENCE SCORE}} = 0.24$, $t(9) = 0.28$, $p =$

0.79; script primary items: $M_{\text{DIFFERENCE SCORE}} = 1.07$, $t(9) = 1.19$, $p = 0.26$). Numerical estimates for in-between responses were not affected by memory load (F 's < 1.5).

Coherence effects within category types

As in Experiment 1, we examined the effect of category coherence on multiple category reasoning within the taxonomic primary and script primary item sets. As before, foods within each set were identified as either high or low coherence based on pretest ratings. People examined more categories in information search for script primary foods associated with low coherence categories, than for script primary foods associated with (relatively) high coherence categories, $F(1, 121) = 3.85$, $p = 0.04$, partial $\eta^2 = 0.05$. A similar effect of coherence approached significance for the taxonomic primary items, $F(1, 121) = 3.08$, $p = 0.08$, partial $\eta^2 = 0.03$.² These effects were not influenced by property type or memory load (F 's < 2.5).

Discussion

This study replicated many of the main trends found in Experiment 1. We again found that multiple category reasoning, as measured by the number of categories consulted and likelihood of an in-between prediction, was more common for script primary than for taxonomic primary foods. Once again, the relative coherence of the categories with which an item was associated appeared to be a major determinant of multiple category reasoning. Moreover, we found an additional effect of coherence within script primary items, with multiple category reasoning most common for food items associated with the least coherent of the script categories.

Experiment 1 found a bias towards taxonomic categories in inductive prediction about cross-classified foods. This experiment examined whether this was due to the use of properties that were linked to taxonomic category membership. When participants made predictions about biochemical or blank properties, we again found a bias favoring the use of taxonomic categories. In these property conditions, taxonomic categories were generally examined before script categories, and given more weight in inductive predictions for tax primary items. In contrast, when people had to make predictions about the time or

² This weaker effect may reflect the restricted range of informativeness/similarity ratings for taxonomic categories (see Table 2).

context in which a food was eaten, script based categories were more likely to be consulted early during information search and the taxonomic bias in numerical estimates was eliminated.

Another important novel finding of Experiment 2 was that levels of multiple category reasoning were dependent on available processing resources. When concurrent processing load was high, multiple category reasoning was reduced, especially for script primary items. Notably, processing load seemed to primarily impact on the information search component of the induction test. People were less likely to consult multiple relevant categories under high load. Once multiple categories were consulted, however, people in the two load conditions were equally likely to make a multiple category prediction.

General discussion

These studies examined the extent to which information from multiple categories is used in inferences about cross-classified objects. The main finding was that multiple category reasoning depended on whether objects were primarily associated with highly coherent taxonomic categories (taxonomic primary items) or with less coherent script categories (script primary items). For taxonomic primary items people usually used single category reasoning, basing their inductive predictions on information from the taxonomic category alone. Multiple category reasoning was more common when foods were primarily associated with a less coherent script category. For these items people were more likely to consult information from both script and taxonomic alternatives and combine this information when making an inductive prediction.

These findings challenge accounts of cross-classification which assume that activation of one category automatically inhibits consideration of other alternatives (Macrae et al., 1995; Nelson & Miller, 1995). Our data clearly show that, for certain kinds of cross-classified instances, people consult more than one relevant category when making inductive predictions. Previous work (e.g., Patalano et al., 2006) has shown that people prefer to base their inductive predictions on more as opposed to less coherent categories. Our studies extend this work in an important and novel way by showing that when less coherent category alternatives are activated first, then people often defer their predictions about a cross-classified item until they have consulted other relevant categories.

We interpret the differences in multiple category reasoning between tax primary and script primary items

as reflecting the relative coherence of taxonomic and script alternatives. It should be acknowledged, however, the current design does not allow us to disentangle the effects of coherence per se from that of food type. Some of the differences between foods identified as script or taxonomic primary may be directly related to differences in perceived coherence. For example, many script primary foods like “kebabs” or “chicken wraps” include ingredients from more than one taxonomic food category, and this may have contributed to their lower level of perceived coherence. Other differences (e.g., script primary foods generally involved more preparation or processing than tax primary foods), however, may not be as strongly correlated with coherence. It is possible, therefore, that such differences may have contributed to the patterns of multiple category reasoning found for tax and script primary items.

While we cannot rule out such effects, it seems highly unlikely that they could supplant relative category coherence as a way of explaining the observed patterns of multiple category reasoning. Recall that there was a large separation between the relative coherence of taxonomic and script categories. The most coherent script primary item received lower ratings than the least coherent tax primary item. This suggests that the coherence difference was a highly salient distinction between tax and script primary items. More generally, in both studies we found direct evidence of an effect of coherence on multiple category reasoning when the confound with food type was removed (i.e., between low and high coherence script primary items).

Experiment 2 examined whether multiple category reasoning for cross-classified items depended on the property being predicted. The property manipulation did not affect the likelihood of examining more than one category or making an in-between prediction. Property type, however, did have subtle effects on the order in which categories were examined and the weighting of information from script and taxonomic categories. In Experiment 1, and in the taxonomic and blank property conditions of Experiment 2, we found a bias towards early examination of taxonomic categories and overweighting of taxonomic information in inductive predictions. This bias was attenuated (and in some cases, eliminated) when participants made predictions about script-relevant situational properties.

These results parallel two trends found in previous work on property effects in category-based induction. The shift towards greater reliance on script-based categories for predictions involving situational properties is consis-

tent with previous findings that the use of different kinds of properties can cause marked shifts in patterns of inductive generalization (e.g., Heit & Rubenstein, 1994; Ross & Murphy, 1999). Our findings are notable in that they extend such property effects to a multiple-category reasoning task.

Our results are also consistent with previous findings of a general preference for reasoning based on taxonomy over other kinds of relations in folkbiological inference (e.g., Ross & Murphy, 1999; Shafto et al., 2007). In particular, the finding of a taxonomic bias for blank properties suggests that, when little is known about a property, taxonomic relations are often used as a default guide for inductive prediction (cf. Shafto & Coley, 2003).

Another important finding was that consideration of multiple categories was reduced for script primary items when there was an increase in concurrent processing demands. Notably, these load effects were limited to information search; concurrent memory load did not impact on whether people made numerical predictions based on both relevant categories, nor did it affect the weighting of information from each category. This suggests that multiple category reasoning involves explicit or intentional search of memory for relevant category information rather than automatic activation of these details (cf. Newell, Paton, Hayes, & Griffiths, *in press*; Verde, Murphy, & Ross, 2005). These results parallel those of studies where property inferences could be based on different types of relations between base and target categories. Shafto et al. (2007) told participants about a novel gene or a disease that was true of one category of animals and asked them to rate the likelihood that taxonomically or ecologically related animals had the same property, under speeded or unspeeded conditions. Differences in the pattern of property generalization were observed for taxonomically related species independent of time pressure, but were only observed for ecologically related species in the unspeeded condition. Hence, retrieval of more complex category relations depends on the availability of sufficient processing resources.³

In this study we focused on the food domain but our results are likely to generalize to other cross-classified

objects. Although most objects can be cross-classified they are likely to vary in the extent to which they are seen as typical of either taxonomic categories or more context dependent or relational categories. For example, a dolphin is both an aquatic animal and a mammal but is more likely to be seen as a typical member of only the first of these categories. Our data suggest that inductive predictions about this kind of cross-classified instance are more likely to involve consideration of multiple categories than predictions about animals that are more strongly associated with a taxonomic grouping.

Another important question is whether taxonomic categories will be seen as the most coherent base for predictions about cross-classified objects, when compared with other kinds of relational categories involving specific goals (e.g., an “apple” is a fruit but also belongs to the goal-directed category, “an item to take on a picnic”). We suspect that the answer depends on how coherence is defined and measured. Goal-directed categories are often highly informative for judgments based on the central relation that defines the category (e.g., “If apples are unavailable what other sorts of foods should be taken on a picnic?”). But when considered in terms of overall within-category similarity and informativeness about novel features, the evidence suggests that taxonomic groupings are viewed as more coherent than a variety of relational alternatives (Gentner & Kurtz, 2005).

Conclusions

Cross-classified objects are ubiquitous so it is important to discover how we coordinate information from multiple categories when making inferences about them. These studies suggest that the view that we focus on just a single category when making feature predictions about cross-classified items is, at best, an oversimplification. While many people in our studies engaged in single category reasoning, we also found robust evidence for multiple category reasoning. Most notably, multiple category reasoning about cross-classified items was often observed when an item had a strong association with a less coherent category and when adequate cognitive resources were available.

Acknowledgments This work was supported by Australian Research Council Discovery Grant DP0770292 to the first and third authors. We would like to thank Brooke Hahn and Helen Paton for their assistance with data collection. We thank Oren Griffiths, Brad Love and three anonymous reviewers for comments on earlier drafts.

³ Another possibility is that the spatial nature of the dot memory task interfered with locating the relevant categories on screen. If so then those in the high load condition should have been slower to examine the category alternatives. However, response latencies (from item onset until the last category was clicked) were unaffected by memory load regardless of the number of categories examined (all p 's > 0.2).

Appendix

Table 3 Biochemical properties (Experiments 1 and 2) and situational properties (Experiment 2) for which predictions were required

Item	Biochemical (“What percentage of the time is it/does it...?”)	Situational (“What percentage of the time is it/does it...?”)
Apple	High in monosaccharides	Used as part of the Solas initiation ceremony in South America
Bacon	Contain glutathione	Eaten on holidays in Belgium?
Baked carrots	High in epicatechin	Found in children's school lunch boxes in Norway
Celery stick	Source of polyphenol antioxidants	Regarded as a good food to take on a trip to the beach
Cheese	Had its colour modified by benzyl peroxide	Served at the Colours of Ostrava festival in eastern Europe
Corn	High in saponins	Served at the Caribana Festival in Cuba
Grapes	Contain carbonic anhydrase	Rated as a good food to eat for losing weight
Lamb	Source of cholecalciferol	Served on a specials menu in a place called Penrose
Meatballs	Elevate blood triglycerides	It takes less than an hour to prepare
Salami	Neutralise prostaglandins	Sold in university cafeterias
Cornflakes	High in polysaccharides	Rated in a survey as a good food to take on a camping trip
Toast	Dilute folic acid	Eaten in winter in Holland
English muffin	Source of tocopherol	Rated as a good food to eat when you are depressed
Custard	High in disaccharides	Served after sunset during certain religious festivals
Chicken wrap	Contain beta-globulin	Consumed at sports events in Canada
Ice Cream	Contain prostaglandins	Sold in school canteens in Western Australia
Kebab	Contain beta-carotene	Sold by street vendors in the Middle East
Pide	Over stimulate the large intestine	It is served with plastic cutlery
Potato	Support the synthesis of blood cells	Eaten outdoors in France
Sandwich	Aid methionine synthesis	Viewed as a good food to take on road trips

References

- Barr, R. A., & Caplan, L. J. (1987). Category representations and their implications for category structure. *Memory & Cognition*, *15*, 397–418.
- De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science*, *17*, 428–433. doi:10.1111/j.1467-9280.2006.01723.x.
- Gentner, D., & Kurtz, K. J. (2005). Relational categories. In W. K. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. W. Wolff (Eds.), *Categorization inside and outside the lab* (pp. 151–175). Washington, DC: American Psychological Association.
- Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *The British Journal of Social Psychology*, *39*, 113–127. doi:10.1348/014466600164363.
- Hayes, B. K., & Newell, B. R. (2009). Induction with uncertain categories: When do people consider the category alternatives? *Memory & Cognition*, *37*, 730–743. doi:10.3758/MC.37.6.730.
- Heit, E., & Rubinstein, J. (1994). Similarity and property effects in inductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 411–422. doi:10.1037/0278-7393.20.2.411.
- Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1995). The dissection of selection in person perception: Inhibitory processes in social stereotyping. *Journal of Personality and Social Psychology*, *69*, 397–407. doi:10.1037/0022-3514.69.3.397.
- Mervis, C. B., Catlin, J., & Rosch, E. (1976). Relationships among goodness-of-example, category norms, and word frequency. *Bulletin of the Psychonomic Society*, *7*, 283–284.
- Murphy, G. L., & Ross, B. H. (1999). Induction with cross-classified categories. *Memory & Cognition*, *27*, 1024–1041.
- Murphy, G. L., & Ross, B. H. (2007). Use of single or multiple categories in category-based induction. In A. Feeney & E. Heit (Eds.), *Inductive reasoning: Experimental, developmental and computational approaches* (pp. 205–225). New York: Cambridge University Press.
- Nelson, L. J., & Miller, D. T. (1995). The distinctiveness effect in social categorization: You are what makes you unusual. *Psychological Science*, *6*, 246–249. doi:10.1111/j.1467-9280.1995.tb00600.x.
- Newell, B. R., Paton, H., Hayes, B. K., & Griffiths, O. (2010). Speeded induction under uncertainty: the influence of multiple categories and feature conjunctions. *Psychonomic Bulletin and Review*, *16* (6)
- Nguyen, S. P., & Murphy, G. L. (2003). An apple is more than just a fruit: Cross-classification in children's concepts. *Child Development*, *74*, 1783–1806. doi:10.1046/j.1467-8624.2003.00638.x.
- Patalano, A. L., & Ross, B. H. (2007). The role of category coherence in experience-based prediction. *Psychonomic Bulletin & Review*, *14*, 629–634.
- Patalano, A. L., Chin-Parker, S., & Ross, B. H. (2006). The importance of being coherent: Category coherence, cross-classification, and reasoning. *Journal of Memory and Language*, *54*, 407–424. doi:10.1016/j.jml.2005.10.005.
- Rehder, B. (2009). Causal-based property generalization. *Cognitive Science*, *33*, 301–343.

- Ross, B. H., & Murphy, G. L. (1996). Category-based predictions: Influence of uncertainty and feature conjunctions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 736–753. doi:[H10.1037/0278-7393.22.3.736](https://doi.org/10.1037/0278-7393.22.3.736).
- Ross, B. H., & Murphy, G. L. (1999). Food for thought: Cross-classification and category organization in a complex real-world domain. *Cognitive Psychology*, *38*, 495–553. doi:[H10.1006/cogp.1998.0712](https://doi.org/10.1006/cogp.1998.0712).
- Shafto, P., & Coley, J. D. (2003). Development of categorization and reasoning in the natural world: Novices to experts, naive similarity to ecological knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 641–649. doi:[10.1037/0278-7393.29.4.641](https://doi.org/10.1037/0278-7393.29.4.641).
- Shafto, P., Coley, J. D., & Baldwin, D. (2007). Effects of time pressure on context-sensitive property induction. *Psychonomic Bulletin & Review*, *14*, 890–894.
- Verde, M. F., Murphy, G. L., & Ross, B. H. (2005). Influence of multiple categories on the prediction of unknown properties. *Memory & Cognition*, *33*, 479–487.