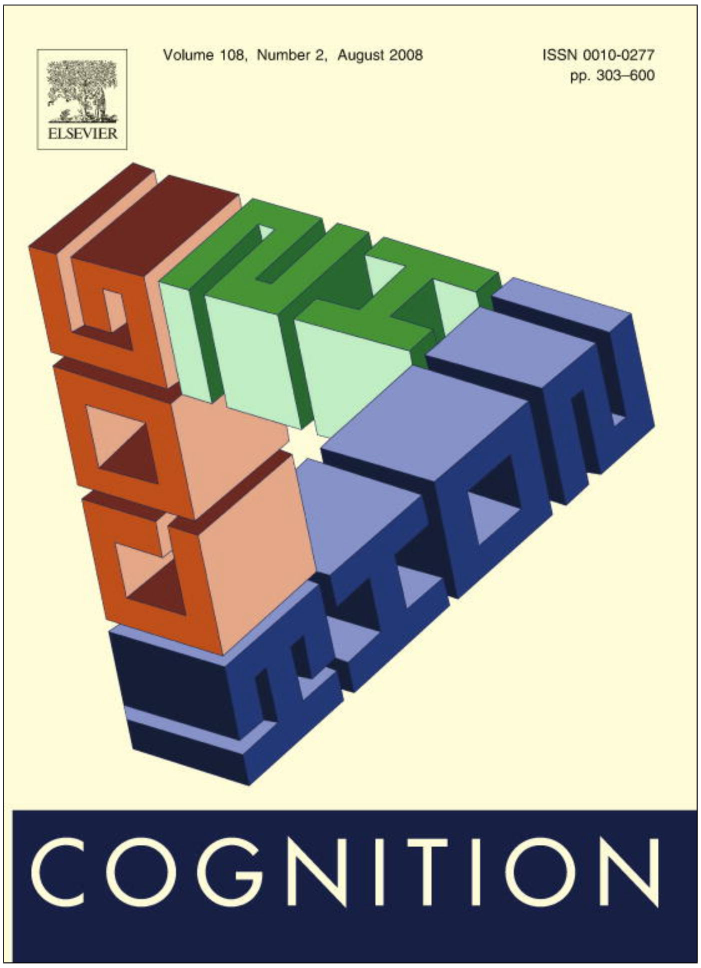
Volume 108, Number 2, August 2008
ISSN 0010-0277
pp. 303–600
ELSEVIER
COGNITION

Available online at www.sciencedirect.com
ScienceDirect
Cognition 108 (2008) 500–506
COGNITION
www.elsevier.com/locate/COGNIT

Discussion

Recognition memory and mechanisms of induction: Comment on Wilburn and Feeney

Vladimir M. Sloutsky *

Center for Cognitive Science, 208C Ohio Stadium East, 1961 Tuttle Park Place, The Ohio State University, Columbus, OH 43210, USA

Received 11 May 2007; revised 29 January 2008; accepted 18 April 2008

Inductive generalization is a critical aspect of human cognition because it enables people to generate new knowledge; however, mechanisms underlying early induction are highly controversial. Some believe that when performing inductive generalizations, people, including young children, first identify the category of an entity and then generalize properties of the entity to other members of the category (see Murphy, 2002, for a review). Therefore, even early in development induction is said to be based on prior categorization of presented entities, and is thus category-based. At the same time, others believe that early induction is based on similarity of compared entities (Sloutsky & Fisher, 2004a). To examine mechanisms of induction across development, Sloutsky and Fisher (hereafter S&F) developed the Induction-then-Recognition task (Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004a, 2004b). S&F argued that if induction is performed on the basis of perceptual information, participants should encode each individual item, whereas if induction is performed on the basis of category-level information, participants may encode only category-level, but not item-specific information. As a result, similarity-based induction should not decrease recognition memory for individual items compared to a baseline condition (in which induction cannot be performed on the basis of categorization), whereas category-based induction should decrease recognition memory compared to

* Tel.: +1 614 688 5855; fax: +1 614 292 0321.
E-mail address: sloutsky.1@osu.edu.

0010-0277/$ - see front matter
doi:10.1016/j.cognition.2008.04.003

the baseline. Therefore, under some conditions (i.e., when baseline accuracy is sufficiently high) similarity-based induction may result in more accurate recognition than category-based induction. These predictions have been confirmed in a number of studies with young children and adults: in contrast to adults, young children (who presumably perform similarity-based induction) exhibited no drop in recognition accuracy compared to the baseline, thus exhibiting greater recognition memory than adults (who presumably perform category-based induction).

Wilburn and Feeney (W&F) re-examine the S&F hypothesis linking mechanism of inductive generalization and recognition memory, with the goal of testing a possible alternative explanation of S&F findings. W&F argue that it is possible that both children and adults perform category-based induction, with children attending more closely to pictures. This elevated attention to pictures results in accurate recognition of individual items.

There are several strengths to W&F's argument. First, W&F offer a hypothesis that is a reasonable alternative to S&F (although a variant of this hypothesis was considered by S&F). They also nicely demonstrate how tricky it could be to distinguish between competing theoretical positions, or even competing hypotheses. W&F conclude that more research is needed to distinguish between the existing theories, and it is difficult to disagree with this general conclusion.

In this response, I address two issues. First, I discuss whether S&F hypothesis is critically undermined by W&F data. And more broadly, I discuss ways of comparing competing hypotheses. In particular, I argue that given two competing hypotheses (Hypothesis 1 and Hypothesis 2), preference should be given to the one that could account for more data. The argument is based on the Bayesian approach to hypothesis testing. If there are several competing hypotheses and a particular body of data, the probability of each hypothesis given the data (H|D) could be compared using the following reasoning. First, given that an *a priori* probability of each hypothesis is unknown, it is not unreasonable to assume that they are equally likely. Therefore, given the same body of data, the hypotheses differ to the extent of differences in their likelihood, or the probability of data given a hypothesis (D|H). This reasoning suggests the following principle of comparison: the broader and more diverse the body of evidence explained by a given hypothesis, the higher the likelihood and therefore the stronger the hypothesis. Therefore, the discussion here may be of interest not only to the students of induction, but also to a broader community interested in how to systematically compare competing hypotheses.

In what follows, I first review W&F's findings. I then compare the likelihood of both hypotheses by comparing how W&F's and S&F's hypotheses account for data stemming from the Induction-then-Recognition paradigm and from some other paradigms.

1. An overview of W&F findings

In Experiment 1, W&F replicate S&F, demonstrating that after performing induction, young children, but not adults, accurately remember pictures that were presented during the induction phase. In Experiment 2, stimuli were presented for

250 ms, and under this condition, children and adults both exhibited high induction accuracy and low recognition accuracy. Furthermore, under the shortened presentation condition, young children exhibited a high gist component (i.e., they readily discriminated a studied category from a novel category) and a low verbatim (or item-specific) component (i.e., they poorly discriminated studied and non-studied members of a studied category). W&F concluded that S&F do not distinguish between the two hypotheses, one proposed by S&F and one proposed by W&F. Recall that Hypothesis 1 (advocated by S&F) is that different mechanisms underlie early and mature induction, with the former being similarity-based and the latter being category-based, whereas Hypothesis 2 (advocated by W&F) is that mechanisms of induction are fundamentally the same, with young children remembering better because they attend to pictures more closely.

2. Comparing the likelihood of the hypotheses

In this section, I review a set of phenomena observed using the Induction-then-Recognition task and consider the ability of each of the two hypotheses to account for each phenomenon. An overview of these phenomena, their sources, and their suggested compatibility with each hypothesis are presented in Table 1. As I argue below, out of the 9 phenomena, there are at least 5 that can be explained by Hypothesis 1, but not Hypothesis 2, there are 3 phenomena that can be explained by both hypotheses, and there is 1 phenomenon that is inconclusive. At the same time, there are no phenomena that can be uniquely explained by Hypothesis 2. Based on this argument I conclude that Hypothesis 1 is stronger than Hypothesis 2, as Hypothesis 2 does not generate evidence that Hypothesis 1 cannot explain. Arguments presented in the table support the central claim of W&F that some of the findings within the Induction-then-Recognition paradigm (e.g., those presented in the first three rows of the table) are compatible with both hypotheses. At the same time, the analysis does not support the conclusion that the Induction-then-Recognition data cannot distinguish between Hypothesis 1 and Hypothesis 2.

Finding 1: Differential memory accuracy in children and adults. Recall that S&F (Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004a, 2004b) have found that after performing induction, children exhibit more accurate recognition memory than adults. In particular, adults confuse studied and novel items of the studied category, whereas children do not. These findings were replicated by Wilburn and Feeney (2008). As argued by W&F (and I agree with their argument), these findings are compatible with either Hypothesis 1 (different mechanisms of induction) or Hypothesis 2 (closer attention to pictures by young children).

Findings 2–3: High induction accuracy coupled with high gist component, and with low recognition accuracy in children under shortened presentation time. Wilburn and Feeney (2008) report two important findings. First, when items were presented for a shortened period of time, young children readily distinguished between items coming from the studied category and from a new category (i.e., there was a high gist component). And second, under the shortened presentation conditions, children

Table 1
Reported findings, their sources, and compatibility with hypotheses proposed by S&F and by W&F

Finding	Source	Compatibility with proposed hypotheses	
		Hypothesis 1	Hypothesis 2
1. After performing induction, children, but not adults exhibit accurate memory	Fisher and Sloutsky (2005); Sloutsky and Fisher (2004a, 2004b); Wilburn and Feeney (2008)	Yes	Yes
2. High gist component in children under shortened presentation time	Wilburn and Feeney (2008)	Yes	Yes
3. High induction accuracy and low recognition accuracy under shorter inspection times	Wilburn and Feeney (2008)	Yes	Yes
4. Differential latencies in children and adults	Wilburn and Feeney (2008); Fisher and Sloutsky (2005); Sloutsky and Fisher (2004a, 2004b)	Although both hypotheses can tentatively account for this finding, finding 4 is inconclusive at this time	
5. Children's recognition accuracy decreases after training	Fisher and Sloutsky (2005); Sloutsky and Fisher (2004a, 2004b)	Yes	No
6. Training affects the induction but not the baseline condition	Fisher and Sloutsky (2005); Sloutsky and Fisher (2004a, 2004b)	Yes	No
7. The differential retention of training across age groups	Fisher and Sloutsky (2005)	Yes	No
8. Gradual decrease in memory accuracy in the course of development	Fisher and Sloutsky (2005)	Yes	No
9. High recognition accuracy for artificial stimuli in adults	Fisher and Sloutsky (2005)	Yes	No

failed to discriminate the studied members of the studied category from new members of the studied category (i.e., there was a low verbatim or item-specific component). Hypothesis 2 has a clear way of accounting for these data – participants perform category-based induction and under shortened presentation they do not have time to fully inspect the pictures. Hypothesis 1 can explain these findings without assuming category-based induction. The underlying idea is that although children perform similarity-based induction, they do not have the time to fully process the items. Given that category members have much in common, these common features are likely to be present on every trial. Therefore, children are more likely to encode the frequently presented common features than distinct features and, as a result, they readily distinguish studied categories from novel categories. At the same time, participants may not have enough time to encode distinct features (these features are important for differentiating between studied and new members of studied categories), and as a result, they poorly discriminate between studied and new members of studied categories. Therefore, Hypothesis 1 and Hypothesis 2 can both account for these findings.

Finding 4: Differential latencies in children and adults. Wilburn and Feeney (2008) report that children and adults exhibited different response latencies, with children having longer latencies than adults. W&F conclude that children were inspecting pictures longer than adults, which resulted in higher memory accuracy. However, it is not clear that the reported latency differences necessarily reflect differences in inspection times (i.e., they can also reflect differences in the decision speed or in the speed of motor responses). Therefore, this finding is inconclusive until inspection times are separated from other variables potentially contributing to response latency. For example, eye-tracking analysis could provide potentially disambiguating information.

Findings 5–6: Children's recognition accuracy decreases after training in the induction condition, but not in the baseline condition. Another set of findings reported by S&F (Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004a, 2004b) pertains to effects of training to perform category-based induction on recognition accuracy. First, training children to perform category-based induction resulted in decreased recognition memory in the induction condition. And second, training had differential effect in the induction condition and in the baseline condition. If recognition accuracy in children is driven by elevated attention to pictures as suggested by Hypothesis 2, it is not clear why training that did not mention the importance or non-importance of pictures would decrease attention to pictures. Again, Hypothesis 1, but not Hypothesis 2, can explain this finding. Moreover, Hypothesis 2 cannot explain the differential effect of training in the induction and the baseline conditions: if children's accuracy stems from elevated attention to pictures then the same training should have similar effects on visual attention and thus result in similar accuracy.

Findings 7–8: The differential retention of training across age groups and gradual decrease in memory accuracy in the course of development. There are two findings reported in Fisher and Sloutsky (2005). First, while training to perform category-based induction decreased recognition accuracy in both 5- and 7-year-olds, two weeks after training, recognition accuracy of 5-year-olds went back to high pre-training levels, whereas effects of training persisted in 7-year-olds. And second, there was a gradual developmental decrease in memory accuracy, with 11-year-olds being less accurate than 7-year-olds, but more accurate than adults. Hypothesis 1 has a clear explanation for these findings – there is a gradual developmental transition from similarity-based induction to category-based induction, and 7-year-olds are at the beginning of this transition, whereas 5-year-olds are not. However, it is not clear how these findings could be explained by Hypothesis 2 – closer attention to pictures in young children. Of course, one could argue that attention to pictures decreases gradually. However, without specifying which aspects of visual attention lead to such gradual decrease, this explanation would be purely ad hoc.

Finding 9: High recognition accuracy for artificial stimuli in adults. Fisher and Sloutsky (2005) report an experiment, in which adults were presented with the Induction-then-Recognition paradigm with novel artificial bug-like stimuli. It was found that adults in the induction condition exhibited recognition accuracy that was comparable to the no-induction baseline and was higher than their recognition accuracy when induction was performed with familiar animals. Again, Hypothesis 1 has a clear explanation of this phenomenon: similarity-based induction is a default and, when categories are novel, people fall back to this default, which results in high recognition accuracy. At the same time, Hypothesis 2 cannot account for this finding because it explains differences in recognition accuracy as stemming from *developmental* differences in allocation of visual attention.

Findings stemming from a different paradigm: Dissociation of category and appearance information in induction. Although we do not include it in the table, it seems useful to consider data stemming from a different paradigm pitting category and appearance information in the course of induction. Sloutsky, Kloos, and Fisher (2007) have recently reported a series of experiments in which 4–5-year-olds performed category learning, categorization, and induction tasks. Stimuli in these tasks were artificial bug-like creatures, and the target category was bound by a deterministic inclusion rule, with all category members sharing the rule. At the same time, many members of the target category looked similar to members of the contrast category (and different from other members of the target category). It was found that while 4–5-year-olds ably learned the categories and categorized correctly throughout the experiment, their induction was based on appearance similarity, and not on category information. Therefore, when appearance and category information are dissociated, induction follows similarity and not category information. This finding was predicted by Hypothesis 1, whereas Hypothesis 2 does not have an obvious way of accounting for it.

In short, W&F (2008) posed two questions: Does S&F evidence unambiguously support Hypothesis 1, indicating that early induction is similarity-based? And should we consider the possibility that children's recognition accuracy is driven by children attending more closely to the pictures (i.e., Hypothesis 2)? They answered "no" to the first question and "yes" to the second. The analysis presented does not dispute these conclusions, while suggesting that at this time Hypothesis 1 is stronger supported by data than Hypothesis 2.

W&F (2008) critique of S&F raises a broader issue of comparing competing hypothesis: can a scientific hypothesis be fundamentally undermined by demonstrating that one or more of the phenomena generated by this hypothesis is compatible with an alternative explanation? In this response I argued the answer requires a systematic comparison of competing hypotheses. One way of doing this is to compare hypotheses on the basis of their "likelihood" or the amount of evidence each hypothesis can account for. According to this criterion, Hypothesis 1 fares better than Hypothesis 2 at this time. One way to generate stronger support for Hypothesis 2 is to predict phenomena that it can uniquely account for. This is a difficult challenge; however, meeting this challenge would constitute a significant contribution to our understanding of mechanisms of inductive generalization.

Acknowledgement

I thank Anna Fisher and Chris Robinson for their helpful comments.

References

Fisher, A. V., & Sloutsky, V. M. (2005). When induction meets memory: Evidence for gradual transition from similarity-based to category-based induction. *Child Development, 76*, 583–597.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Sloutsky, V. M., & Fisher, A. V. (2004a). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General, 133*, 166–188.

Sloutsky, V. M., & Fisher, A. V. (2004b). When development and learning decrease memory: Evidence against category-based induction in children. *Psychological Science, 15*, 553–558.

Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When looks are everything: Appearance similarity versus kind information in early induction. *Psychological Science, 18*, 179–185.

Wilburn, C., & Feeney, A. (2008). Do development and learning really decrease memory? On similarity and category-based induction in adults and children. *Cognition, 106*, 1451–1464.