



Contents lists available at ScienceDirect

# Journal of Experimental Child Psychology

journal homepage: [www.elsevier.com/locate/jecp](http://www.elsevier.com/locate/jecp)



## Commentary

# What can we learn from toddlers about categorical perception of color? Comments on Goldstein, Davidoff, and Roberson

Anna Franklin<sup>a,\*</sup>, Oliver Wright<sup>b</sup>, Ian R.L. Davies<sup>a</sup>

<sup>a</sup>Department of Psychology, University of Surrey, Surrey GU2 5XH, UK

<sup>b</sup>Department of Psychology, Bilkent University, 06800 Bilkent Ankara, Turkey

## ARTICLE INFO

### Article history:

Received 8 July 2008

Revised 13 August 2008

Available online 1 October 2008

### Keywords:

Color perception

Categorical perception

Color terms

Toddlers

## ABSTRACT

We comment on Goldstein, Davidoff, and Roberson's replication and extension (*Journal of Experimental Child Psychology*, 102, 219–238 [2009]) of our study of the effect of toddlers' color term knowledge on their categorical perception (CP) of color (*Journal of Experimental Child Psychology*, 90, 114–141 [2005]). First, we discuss how best to assess color term knowledge when concerned with the effects of language on color CP. A reanalysis of our data indicates that even toddlers who do not know the terms for the relevant focal colors still show CP. Second, we comment on Goldstein and colleagues' finding of blue–purple CP, as we did, but not of blue–green CP in Himba toddlers. We present contrasting data from Wright (unpublished PhD thesis, University of Surrey, 2006) that demonstrates blue–green CP in Himba toddlers. Finally, we discuss the limitations of the approach taken by all of these investigations and discuss theoretical accounts of the origin and nature of color CP.

Crown Copyright © 2008 Published by Elsevier Inc. All rights reserved.

## Introduction

Categorical perception (CP) of color is shown when discrimination between two colors from different categories is better than discrimination between two equivalently separated colors from the same category (Bornstein & Korda, 1984). The contribution of language to this effect has been debated extensively, with some arguing that CP is dependent on color language (e.g., Roberson, Davies,

\* Corresponding author.

E-mail address: [a.franklin@surrey.ac.uk](mailto:a.franklin@surrey.ac.uk) (A. Franklin).

& Davidoff, 2000) and others arguing that CP can exist in the absence of color language (e.g., Bornstein, Kessen, & Weiskopf, 1976; Franklin & Davies, 2004). One approach we have taken when investigating the role of language in color CP is to test toddlers who differ in the extent of color term acquisition (Franklin, Clifford, Williamson, & Davies, 2005a). In that investigation, we found CP across blue–green, blue–purple, and pink–red boundaries irrespective of whether toddlers could name the stimuli used in the color CP task. A second experiment found blue–purple CP in toddlers who were from the Namibian Himba population. The Himba language has five basic color terms (vs. 11 in English), and the toddlers tested had not yet learned these basic color terms. We concluded that color term knowledge does not affect the extent of color CP in toddlers. In this issue of the *Journal of Experimental Child Psychology*, Goldstein, Davidoff, and Roberson (2009) largely replicate these findings. However, a further analysis suggesting that color CP is dependent on color term knowledge for the wider category, and the absence of blue–green CP in Himba toddlers, leads them to very different conclusions.

In this commentary, we reflect on Goldstein and colleagues' (2009) replication and extension of our study. Their study raises an important question about how to assess color term knowledge when investigating potential effects of language on CP, and we discuss this issue first. We then consider their data on color CP in toddlers from the Himba population. The limitations of the approach taken in both investigations are then discussed. Finally, we discuss the theoretical framework provided by Goldstein and colleagues and offer alternative plausible theoretical accounts.

### Issues of color term knowledge

Goldstein and colleagues' (2009) study raises an important question about how to assess color language when investigating the effects of language on color CP. Is it best to assess naming of the stimuli used in the color CP task, or is it best to assess naming and comprehension for the category as a whole? If the effects of language are due to the online effects of verbal labeling (e.g., explicit labeling of stimuli during the task), naming of the stimuli used in the CP task should be most relevant. In our original study, color CP was unaffected by the pattern of naming for the stimuli used in the color CP task. We see that Goldstein and colleagues' data largely replicate this finding. Those toddlers who named the CP task stimuli inappropriately by naming the same category color (within-category) with different names, naming the different category colors (between-category) with the same name, or naming all colors differently showed blue–purple CP to the same extent as those who named the colors according to the intended categorical membership. There was a trend ( $p = 0.077$ ) for a different pattern of results for the blue–green boundary. As argued by Goldstein and colleagues, these findings could imply that it is not the use of online verbal labeling that produces color CP.

However, language effects might not be due to verbal labeling but rather could arise from some other mechanism. For example, language may warp perceptual space by drawing attention to the category boundary, leading to greater perceptual learning for stimuli around the category boundary than for stimuli within a category. Alternatively, mapping the same word to a range of colors may induce some kind of “category code” or conceptual representation of the category that then, through some currently unspecified mechanism, influences perceptual judgments. If the effects of language are not via verbal labeling, accurate assessment of color term knowledge for the whole category, and not just the stimuli used in the CP task, could be important. Goldstein and colleagues (2009) also use a measure that takes note of the number of overextensions of the term (Soja, 1994). If more than one overextension was made, it was taken that the toddler did not know the term. Therefore, this method assesses color term knowledge for the wider category (see also Pitchford & Mullen, 2002). Goldstein and colleagues find no CP for those who “don't know” the relevant color terms when assessed by Soja's criteria.

To see whether we could provide converging evidence that CP depends on knowledge of the relevant color terms, we reanalyzed the data from our original investigation (Franklin et al., 2005a). If CP depends on knowledge of the relevant color terms, those who do not know the terms for stimuli that are the best examples (focal colors) of the relevant categories should be less likely to show CP than those who know these terms. In our original investigation, naming and comprehension of the 11 basic

focal colors was assessed to see whether general color term fluency correlated with the extent of color CP.<sup>1</sup> In our reanalysis of the data, we identified eight toddlers (mean age = 31 months), who were tested for blue–green or blue–purple CP, who could not name or point to the focal stimuli for both of the relevant categories. These toddlers also occasionally used blue, green, or purple terms inaccurately when naming other focal stimuli, or the toddlers pointed to blue, green, or purple stimuli incorrectly when asked to identify other colors (indicating an awareness of the term but erroneous application). These eight toddlers still showed a significant CP effect,  $t(7) = 2.55, p < .05, d = 1.1$ , with within-category performance ( $M = 46.87\%$ ,  $SD = 33.9$ ) at chance,  $t(7) = 0.26, p = .80$ , and between-category performance ( $M = 75.00\%$ ,  $SD = 18.9$ ) significantly above chance,  $t(7) = 3.74, p < .01$ .

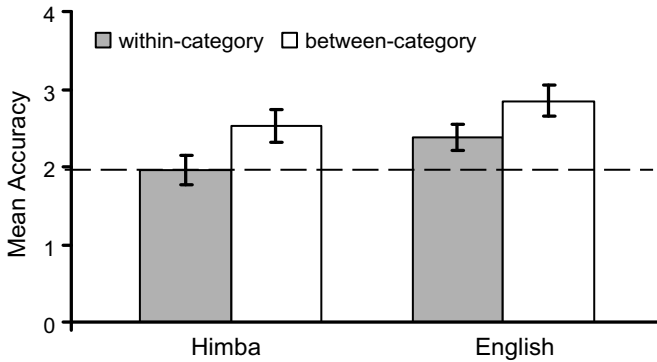
How does this reanalysis compare with Goldstein and colleagues' (2009) analysis using Soja's (1994) criteria? Although comprehension of nonfocal colors was not assessed in our investigation, toddlers would be classified as not knowing color terms according to Soja's criteria because the toddlers in our reanalysis could not even name or point to the relevant focal stimuli. Therefore, our reanalysis contrasts with Goldstein and colleagues' findings by showing that, at least for our data, CP occurs even in those who do not know the names of the category foci. We return to further discussion of these contrasting findings later in this commentary.

### Color CP in Himba toddlers

Goldstein and colleagues (2009) provide much needed further data on color CP in young children at the early stages of color term acquisition from populations whose language has fewer basic color terms than English. As in our original investigation, Goldstein and colleagues' data suggest that there is CP across the blue–purple boundary in Himba toddlers. Whereas we suggested that the presence of blue–purple CP in Himba toddlers indicates a nonlinguistic category effect, Goldstein and colleagues attribute the effect to language. One explanation offered by them is that although toddlers seemed not to know color terms on testing, they did actually have some knowledge of the Himba terms *zoozu* (black/dark) and *burou* (blue or green). In previous research, these terms were sometimes offered by Himba adults when naming colors within the range of the stimuli used in the blue–purple color CP task. This argument rests on the assumption that language can induce color CP even when color terms are not known well enough to be provided on naming and comprehension tasks. If further investigation supports this account, it would certainly provide evidence for the powerful effects of language on perception. Goldstein and colleagues' second explanation for the blue–purple CP in toddlers is that the toddlers may have had some knowledge of the cattle terms that Himba adults sometimes use for purple. This argument rests on three assumptions: (a) that Himba toddlers are exposed to the cattle naming terms in question and that these terms are part of their vocabulary, (b) that Himba toddlers would apply them to color and not exclusively to cattle, and (c) that the degree of exposure and use would be consistent and sufficient enough to induce a category effect at the blue–purple boundary. Again, if further investigation supports these assumptions, it would provide evidence for the strong effects of language on perception in very young children.

Testing for blue–green CP in Himba toddlers provides a safer test of CP in young children with no relevant color term knowledge because Himba have a single basic term, *burou*, covering blue and green. Here Goldstein and colleagues (2009) do not find CP. We were surprised by this result because data that one of us (Wright) had collected during doctoral research indicated a clear blue–green CP effect for Himba toddlers. Wright (2006) tested for blue–green CP in 25 Himba 2- to 4-year-olds (data from 33 Himba are analyzed by Goldstein et al.), as well as 26 English 2- to 4-year-olds. The same pro-

<sup>1</sup> Franklin and colleagues (2005a) reasoned that if language affected CP via verbal labeling, general naming ability (as indicated by number of focal terms known) may be associated with the size of the category effect. For example, a child who knows all of the focal terms may be more likely to use verbal labels on the CP task than a child who knows fewer focal terms due to greater proficiency in color naming. However, Franklin and colleagues found a very small negative relationship between number of focal terms known, and Goldstein and colleagues (2009) found no significant relationship.



**Fig. 1.** Mean accuracy ( $\pm$ SE) for within- and between-category trials for the blue–green boundary for Himba and English 2- to 4-year-olds. The dashed line indicates performance at chance. Data for the figure were taken from Wright (2006).

cedure was used as in Experiment 2 of Franklin et al. (2005a), with the blue–green stimuli from Experiment 1 (as in Goldstein et al.).<sup>2</sup> Fig. 1 gives the mean accuracy for within- and between-category conditions for Himba and English toddlers.

As can be seen in Fig. 1, there was an effect of category,  $F(1, 49) = 7.91, p < .05, \eta_p^2 = 0.13$ , and no category by language group interaction,  $F(1, 49) = 0.07, p = 0.80, \eta_p^2 = 0.13$ . This finding clearly contrasts with Goldstein and colleagues' (2009) lack of blue–green CP in the Himba toddlers the authors tested.

### Limitations of the approach

Testing toddlers to study the effects of language on CP has yielded contrasting results. Data from our original investigation (Franklin et al., 2005a), and to a large extent Goldstein and colleagues' (2009) data, suggest that the pattern of naming of the stimuli used in the CP task does not affect the extent of CP. Goldstein and colleagues' study suggests that color term knowledge for the wider category is needed for color CP, whereas a reanalysis of our data suggests that those who do not know the term for the best examples of the relevant categories still show a category effect. Both of our data suggest that blue–purple CP is shown by Himba toddlers who appear to have no knowledge of color terms and whose language has fewer basic color terms than English. However, Goldstein and colleagues attribute this to language. Finally, although Goldstein and colleagues do not find blue–green CP in Himba toddlers, Wright (2006) clearly did find a category effect. Clearly, no consensus about the role of language in color CP has emerged from these studies.

It may, at this stage, be useful to reflect on the limitations of the approach taken in these toddler investigations. One limitation is the difficulty of proving a negative—establishing that a toddler has no knowledge of a color term at all. As discussed earlier, toddlers might not provide the accurate term during formal color comprehension and naming tasks, yet toddlers may have been exposed to the term and may indeed have some underlying knowledge of the term or some level of competence with the term. The CP shown by toddlers with no explicit knowledge of the terms for the relevant focal colors in our reanalysis of our data could be explained by arguing that these children did indeed have some knowledge of the relevant terms (some competence) but were just not yet able to use them on formal comprehension and naming tasks (a performance limitation). This limitation is less of an issue if the proposed mechanism of the influence of language is via verbal labeling; if toddlers could use the label on the CP task, they would surely be able to identify the color on a comprehension task. However, if the proposal is that general exposure to color terms warps perceptual color space, being able to identify latent knowledge of the term could be important. The only clear way to establish whether CP exists in toddlers in the absence of the relevant color language is to test toddlers whose

<sup>2</sup> All methodological details of the experiment can be found in Franklin et al. (2005a). However, a copy of the writeup of the experiment in Wright's thesis can also be made available on request.

language does not mark the distinction (e.g., testing for blue–green CP with the Himba toddlers). However, given the contrasting findings of Goldstein et al. (2009) and Wright (2006) on this issue, even this does not resolve the question, and further research is clearly needed.

We suggest that this further research should use different tasks from those used with toddlers so far. Various limitations of the color memory task are apparent. First, from Goldstein and colleagues' (2009) data, it is evident that 2- to 4-year-olds are prone to perform at ceiling or chance (approximately half of the children were at ceiling or chance in their experiments). It appears that the task, at least with the parameters set as they are, might not be suitable for the age range in which we are interested. Second, although time to respond could be a useful additional measure, it cannot be recorded accurately enough to be useful with the current version of the task. Third, when we decided to use this task, testing for CP with memory tasks was standard (e.g., Roberson & Davidoff, 2000). However, since that time, adult studies have used tasks that remove the memory component such as visual search (e.g., Daoutis, Pilling, & Davies, 2006; Gilbert, Regier, Kay, & Ivry, 2006; Roberson, Pak, & Hanley, 2008) and target detection tasks (Drivonikou et al., 2007; Franklin et al., 2008; Franklin, Pilling, & Davies, 2005b). When these tasks are used, it is clearer that any categorical effects are actually related to perceptual processes rather than to memory effects such as the shift toward prototype (Huttenlocher, Hedges, & Vevea, 2000). Therefore, we suggest that future CP studies, if testing toddlers, should use tasks such as visual search and target detection tasks that are adapted for use with toddlers. Such studies should also try to use large samples so that if they are split into subgroups, statistical comparisons still have sufficient power. This is especially important when findings rest on the lack of a significant effect such as for the “don't know” group in Goldstein and colleagues where there is no significant CP but only eight participants, meaning that the lack of an effect could be attributed to lack of statistical power.

However, it might also be necessary to review this approach altogether. What is it that we are trying to find out by testing for color CP in toddlers at the stage of color term acquisition? Our aim in the original investigation was to see whether, as the verbal labeling account of CP would predict, being able to name colors used in a CP task affects the extent of CP. There is now converging and convincing evidence that language does affect CP in adults; for example, recent evidence suggests that CP is left hemisphere dominant in adults (Drivonikou et al., 2007; Franklin et al., 2008; Gilbert et al., 2006; Roberson et al., 2008). However, just because CP appears to be mediated by language in adults, it does not follow that language is the origin of the effect. Categories present early on in development may indeed become attenuated if language does not support the distinction (McDonough, Choi, & Mandler, 2003). Therefore, the important question is whether color CP exists before language acquisition.

As discussed earlier, toddlers, even if they fail on comprehension and naming tasks, may have implicit knowledge of color terms (or even cattle terms). Therefore, to address the question of whether CP exists before language acquisition, a cleaner approach is to test prelinguistic infants. Here the evidence suggests that color CP can exist in the absence of language (Bornstein et al., 1976; Franklin & Davies, 2004; Franklin et al., 2008; Franklin et al., 2005b). If familiarized to a given hue, 4-month-olds will show novelty preference only when novel and familiar hues are between-category rather than within-category (e.g., Franklin & Davies, 2004). In addition, infants will fixate or initiate an eye movement to a chromatic target on a chromatic background faster when the target and background are between-category than when they are within-category (Franklin, Pilling et al., 2005b; Franklin et al., 2008). These infant CP effects are found using the standard approach to testing for CP, that is, when within- and between-category stimuli are equated in Munsell color space or the CIE 1976 ( $L^* u^* v^*$ ) color metric. The only published study not to find a category effect in infants (Gerhardstein, Renner, & Rovee-Collier, 1999) had an unintended category boundary in one of the within-category pairs and the use of incandescent lighting warped the color metric, rendering the conclusions unsafe (Davies & Franklin, 2002). As Goldstein and colleagues (2009) highlight, the challenge is to reconcile the evidence for infant color CP with the evidence showing that CP is dependent on language. We elaborate on this and the wider theoretical argument below.

## Theoretical accounts

Goldstein and colleagues (2009) frame the debate about the origin and nature of color CP in terms of Nativist and Relativist accounts—the nativist account that color space is ‘warped only at the bound-

aries between a putative set of hardwired and innate color categories' and the relativist account that CP is dependent on language. As others have argued (Kay & Regier, 2006), framing the debate in terms of nativist and relativist accounts might not be particularly helpful. It is clear that there is evidence for both language effects on CP and evidence for prelinguistic color CP in infants. The "either/or" Nativist/Relativist debate clearly does not provide an adequate framework for all findings in the field. However, there is evidence that prelinguistic color CP and language-mediated CP can exist alongside each other. Recent research suggests that color CP is lateralized to the right hemisphere in infants rather than to the left hemisphere as in adults (Franklin et al., 2008). The effect may suggest that there are two forms of CP: one that is nonlexicalized and lateralized to the right hemisphere (as in infants) and one that is lexicalized and lateralized to the left hemisphere (as in adults). We are currently conducting further research to test the robustness of the effect, to test whether the change in hemisphere is related to color term acquisition, and to clarify the relationship and interaction between right and left hemisphere forms of CP. If further testing supports this hypothesis, it could help to resolve the debate.

Although infants as young as 4 months certainly show color CP, it does not necessarily follow that it is hardwired into the human visual system. It is entirely possible that category effects are acquired. Processes that have been proposed to account for the origin of color categories in language, such as sensitivity to statistical tendencies in the color of surfaces in our environment (Yendrikhovskij, 2001) or properties of ambient light sources (Shepard, 1992), could induce infant color category effects. For example, regional differences in the ambient illumination during infancy can lead to long-term differences in color vision (Laeng et al., 2007). There is also converging evidence that infants can acquire categories rapidly on the basis of statistical tendencies and correlations among stimuli (e.g., Younger, 1990). Of course, if acquired, color CP in infants need not necessarily be universal, yet acquired categories could be universal if the conditions for acquisition were shared across cultures. These issues also await further testing, and cross-cultural investigations of infant color CP may be particularly useful in determining the origins of the effect.

## Conclusions

Goldstein and colleagues (2009) largely replicate our investigation of the effect of color term knowledge on CP of color in toddlers (Franklin et al., 2005a). However, an analysis of color CP according to toddlers' color term knowledge for the wider category, and the absence of blue–green CP in Himba toddlers, lead Goldstein and colleagues to very different conclusions. Although their data provide some evidence for the absence of color CP when color terms are not known, this commentary has highlighted contrasting evidence. A reanalysis of Franklin and colleagues' (2005a) original data indicated that even toddlers who did not know the terms for the best example of the relevant categories showed CP, and data from Wright's (2006) doctoral thesis provided evidence for blue–green CP in Himba toddlers whose language does not make a blue–green distinction. Therefore, it appears that currently it is difficult to form a clear conclusion from investigations of CP in toddlers about the origin and nature of color CP. We emphasize the importance of further research that overcomes the limitations of the current approach and may provide a more consistent account of the origin and nature of CP of color.

## Acknowledgments

We thank David Bjorklund and an anonymous reviewer for their critical reading and helpful comments on a previous draft of the manuscript.

## References

- Bornstein, M. H., Kessen, W., & Weiskopf, S. (1976). Color vision and hue categorization in young infants. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 115–129.
- Bornstein, M. H., & Korda, N. O. (1984). Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing. *Psychological Research*, 46, 207–222.
- Daoutis, C. A., Pilling, M., & Davies, I. R. L. (2006). Categorical effects in visual search for colour. *Visual Cognition*, 14, 217–240.

- Davies, I. R. L., & Franklin, A. (2002). Categorical similarity may affect color pop-out in infants after all. *British Journal of Developmental Psychology*, *20*, 185–203.
- Drivonikou, G. V., Kay, P., Regier, T., Ivry, R. B., Gilbert, A. L., Franklin, A., & Davies, I. R. L. (2007). Further evidence for lateralization of Whorfian effects to the right visual field. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 1097–1102.
- Franklin, A., Clifford, A., Williamson, E., & Davies, I. R. L. (2005a). Color term knowledge does not affect categorical perception of color in toddlers. *Journal of Experimental Child Psychology*, *90*, 114–141.
- Franklin, A., & Davies, I. R. L. (2004). New evidence for infant colour categories. *British Journal of Developmental Psychology*, *22*, 349–377.
- Franklin, A., Drivonikou, G. V., Bevis, L., Davies, I. R. L., Kay, P., & Regier, T. (2008). Categorical perception of color is lateralized to the right hemisphere in infants, but to the left hemisphere in adults. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 3221–3225.
- Franklin, A., Pilling, M., & Davies, I. R. L. (2005b). The nature of infant color categorisation: Evidence from eye-movements on a target detection task. *Journal of Experimental Child Psychology*, *91*, 227–248.
- Gerhardstein, P., Renner, P., & Rovee-Collier, C. (1999). The roles of perceptual and categorical similarity in colour pop-out in infants. *British Journal of Developmental Psychology*, *17*, 403–420.
- Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 489–494.
- Goldstein, J., Davidoff, J., & Roberson, D. (2009). Knowing color terms enhances recognition: Further evidence from English and Himba. *Journal of Experimental Child Psychology*, *102*, 219–238.
- Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgement? *Journal of Experimental Psychology: General*, *129*, 220–241.
- Kay, P., & Regier, T. (2006). Language, thought, and color: Recent developments. *Trends in Cognitive Sciences*, *10*, 51–54.
- Laeng, B., Brennen, T., Elden, A., Garre Paulsen, H., Banerjee, A., & Lipton, R. (2007). Latitude-of-birth and season-of-birth effects on human color vision in the Arctic. *Vision Research*, *47*, 1595–1607.
- McDonough, L., Choi, S., & Mandler, J. M. (2003). Understanding spatial relations: Flexible infants, lexical adults. *Cognitive Psychology*, *46*, 229–259.
- Pitchford, N. J., & Mullen, K. T. (2002). *Is the acquisition of basic colour terms in young children constrained?* *Perception*, *31*, 1349–1370.
- Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory and Cognition*, *28*, 977–986.
- Roberson, D., Davies, I. R. L., & Davidoff, J. (2000). Color categories are not universal: Replications and new evidence from a Stone Age culture. *Journal of Experimental Psychology: General*, *129*, 369–398.
- Roberson, D., Pak, H., & Hanley, J. R. (2008). Categorical perception of color in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition*, *107*, 752–762.
- Shepard, R. N. (1992). The perceptual organization of colors: An adaptation to the regularities of the terrestrial world. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 494–532). Oxford, UK: Oxford University Press.
- Soja, N. N. (1994). Young children's concept of color and its relation to the acquisition of color words. *Child Development*, *65*, 918–937.
- Wright, O. (2006). The origin and nature of perceptual colour categories. Unpublished PhD thesis, University of Surrey.
- Yendrikhovskij, S. N. (2001). Computing color categories from statistics of natural images. *Journal of Imaging Science and Technology*, *45*, 409–418.
- Younger, B. (1990). Infants' detection of correlations among feature categories. *Child Development*, *61*, 614–620.