Available online at www.sciencedirect.com
ScienceDirect
Cognition 106 (2008) 665–681
COGNITION
www.elsevier.com/locate/COGNIT

Labels can override perceptual categories in early infancy

Kim Plunkett [a,*], Jon-Fan Hu [a,1], Leslie B. Cohen [b,2]

[a] *Oxford University, Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD, UK*
[b] *University of Texas, Austin, USA*

Received 14 October 2006; revised 26 March 2007; accepted 7 April 2007

Abstract

An extensive body of research claims that labels facilitate categorisation, highlight the commonalities between objects and act as invitations to form categories for young infants before their first birthday. While this may indeed be a reasonable claim, we argue that it is not justified by the experiments described in the research. We report on a series of experiments that demonstrate that labels can play a causal role in category formation during infancy. Ten-month-old infants were taught to group computer-displayed, novel cartoon drawings into two categories under tightly controlled experimental conditions. Infants were given the opportunity to learn the two categories under four conditions: Without any labels, with two labels that correlated with category membership, with two labels assigned randomly to objects, and with one label assigned to all objects. Category formation was assessed identically in all conditions using a novelty preference procedure conducted in the absence of any labels. The labelling condition had a decisive impact on the way infants formed categories: When two labels correlated with the visual category information, infants learned two categories, just as if there had been no labels presented. However, uncorrelated labels completely disrupted the formation of any categories. Finally, consistent use of a single label across objects led infants to learn one broad category that included all the objects. These findings demonstrate that even before

[*] Corresponding author. Tel.: +44 01865 271398; fax: +44 01865 310447.
E-mail addresses: kim.plunkett@psy.ox.ac.uk (K. Plunkett), jon-fan.hu@psy.ox.ac.uk (J.-F. Hu), cohen@psy.utexas.edu (L.B. Cohen).
[1] Tel.: +44 01865 271398; fax: +44 01865 310447.
[2] Tel.: +1 512 4710189; fax: +1 512 4715935.

doi:10.1016/j.cognition.2007.04.003

infants start to produce their first words, the labels they hear can override the manner in which they categorise objects.

Keywords: Language acquisition; Categorisation; Infancy; Cognitive development; Novelty preference; Perceptual similarity; Statistical learning; Lexical development

1. Introduction

What role does language play in early infant categorisation processes? Could hearing the word *toy* uttered whilst attending to a toy contribute to the infant's appreciation that all other objects thus labelled belong to one and the same category of objects? Such a capacity would have profound implications for the manner in which a caregiver's labelling practises shaped the conceptual structure of the infant's mind – what counts as a toy? Labels have the potential to offer a receptive infant an auditory window on the conceptual world of adults. One might be tempted to suppose that such a capacity would be operational by the time an infant has mastered an extensive vocabulary and knows what many things should be called. But what effect would labels have on the pre-linguistic infant, devoid of speech and of limited comprehension? Can adult vocabulary penetrate and modulate the foundational structure of the concepts that infants are acquiring during their first year of life?

We will demonstrate in a series of five experiments that the answer to this question is probably *yes*. *Probably* because our experimental procedures are limited to the laboratory and our experimental stimuli are limited in scope. We will argue that a convincing demonstration that labels can impact the process of infant categorisation requires that infants learn *novel* categories which they could not bring with them from outside the laboratory, and that the structure of the novel categories they learn in the laboratory is independently motivated by the labelling contingencies to which the infants are exposed. In other words, infants should not be able to acquire the novel categories without the labels. To realise these strict operational criteria, it is necessary to demonstrate that infants can acquire at least two novel categories in the laboratory. It is not easy to show that infants have learnt two novel categories. Fortunately, this work has been done before us (Younger, 1985). We will build on this work and show that we can determine the structure of the novel categories that infants learn by manipulating the labelling events during category formation. In particular, we will show that by exposing infants to one or two labels we can determine whether they learn one category, two categories or no category at all.

1.1. Background

A substantial body of research over the past decade has laid claim to the view that words highlight the commonalities between objects (Waxman, 1999; Waxman & Booth, 2003), facilitate the categorisation process in infants (Balaban & Waxman, 1997; Fulkerson & Waxman, 2007), and that words serve as invitations to form new categories and concepts (Waxman & Markow, 1995). Using an adaptation of the well-established novelty preference procedure (Fantz, 1964; Roder, Bushnell, & Sasseville, 2000; Wetherford & Cohen, 1973), in which infants are initially familiarised to a sequence of individual objects taken from a single category and then tested on a pair of new objects, one of which is from a novel category, Waxman and colleagues have demonstrated that the presence of a label during the familiarisation phase can influence the ability of infants, as young as 6 months of age, to respond on the basis of an object's category membership. In particular, they show that infants show a novelty preference for out-of-category objects in a labelling condition but not in conditions that do not involve novel words.

Given these claims, and the apparent evidence for them, it is important to note that there are a variety of difficulties in interpreting the causal role of words in these experiments. First, consider the claim that labels serve as 'invitations to form categories': Although, many of the object *tokens* used in the familiarisation phase of these experiments were most likely novel to the infant participants, they were nevertheless taken from categories with which many infants would be familiar through everyday experience, with toys or picture books – cars, planes, horses, dinosaurs, fish, animals, vehicles, tools, fruits, etc. Therefore, infants may have already formed categories of some or all of these objects before they participated in the experiment, reducing the experimental role of labels to one of category *re-activation* rather than category formation. The only clear use of novel objects is described by Booth and Waxman (2002) who tested 14- and 18-month-old infants in a category-matching task. In this study, 14-month-old failed to show an effect of labelling unless they were also provided with cues to the core meaning (function) of the objects. This result indicates that labels themselves may not be the primary locus of the categorisation effect. Prior familiarity with non-linguistic characteristics of the visual stimuli may be a pre-condition for the formative effect of any labels.

Do words facilitate categorisation by young infants? Facilitation needs to be demonstrated against a baseline condition. In many experiments, the baseline comparison is a no-word condition where objects are introduced with a carrier phrase such as "Look what's here" or "See what I have" (Booth & Waxman, 2002, 2003; Waxman, 1999; Waxman & Booth, 2003; Waxman & Markow, 1995). Other studies have reported that infants show out-of-category novelty preferences (e.g., Behl-Chadha, 1996; Eimas & Quinn, 1994; Quinn & Eimas, 1996; Younger, 1985) in the absence of any auditory stimulation during the familiarisation phase of the experiment. Therefore, an important comparison is whether labels enhance out-of-category novelty preferences over and above any such preferences that might be observed in the absence of auditory stimulation. If no differences were observed, then attributing a facilitating effect to words would be unjustified. In the absence of any such comparisons, claims regarding the facilitating effects of words are, at least, premature. Unfortunately, no-sound baseline comparisons are not provided in these studies (Balaban & Waxman, 1997; Booth & Waxman, 2002, 2003; Fulkerson & Waxman, 2007; Waxman, 1999; Waxman & Booth, 2001, 2003; Waxman & Braun, 2005; Waxman & Markow, 1995).

The absence of non-auditory comparisons becomes particularly problematic in the light of recent experimental studies with infants showing that auditory stimuli can *overshadow* the processing of visual stimuli (Robinson & Sloutsky, 2004). Using an habituation task with compound auditory–visual stimuli, Robinson and Sloutsky (2004) have demonstrated that infants fail to detect, i.e., fail to dishabituate to, a change in the visual component of a compound habituation stimulus but they do detect a change in the auditory component. The authors conclude that infants exhibit auditory dominance when processing novel auditory–visual compound stimuli. The absence of a novelty preference effect in the no-word condition ("Look what's here", "See what I have") of Waxman and colleagues' studies may be partly driven by an auditory dominance effect. A similar difficulty arises in a recent study by Fulkerson and Waxman (2007) who demonstrated a novelty preference in 6- and 12-month-old infants in a labelling condition but not in a tone condition. Failure to evaluate performance in a silent condition makes it impossible to determine whether labels are facilitating categorisation or tones are interfering with categorisation.

Some evidence for auditory interference effects is provided by Fulkerson and Haaf (2003) who compared novelty preferences for out-of-category objects in label, non-label and no-sound conditions. When 9- and 15-month-old infants were familiarised with objects from the same basic level category (horses or airplanes), both age groups showed a novelty preference in *all* familiarisation conditions. However, the novelty preference was largest in the no-sound condition, suggesting that the auditory stimuli may have interfered with, rather than facilitated, object categorisation.

It is also worth noting that Fulkerson and Haaf (2003) found that labels produced a novelty preference, whereas the no-sound condition did not when the familiarisation stimuli formed a global level category (animals or vehicles). Although this finding does not accord with other experiments that infants can form global level categories in the absence of auditory stimuli (Behl-Chadha, 1996; Quinn & Johnson, 2000), it does suggest that any facilitatory effect of a label must be highly sensitive to the perceptual similarity of the familiarisation objects. In particular, labels may be particularly efficacious in facilitating categorisation when the familiarisation objects are perceptually dissimilar from each other.

Why does the Word condition produce a novelty preference whereas the no-word condition fails to do so? This is a robust finding across many studies (Booth & Waxman, 2003; Fulkerson & Haaf, 2003; Waxman, 1999; Waxman & Booth, 2003; Waxman & Braun, 2005; Waxman & Markow, 1995). An important clue to the source of this effect is reported by Waxman (1999): During the familiarisation phase of her experiment, 13-month-old infants "devoted more attention to objects when novel words were present". (B42). She concludes that "infants' heightened attention to named objects suggests an attentional mechanism by which novel words may facilitate the acquisition of object categories, particularly at the earliest stages of word learning". (B45) In contrast, the no-word condition in these experiments is unlikely to contain any speech that is novel to the infant.

It is unclear whether the novel words used in these experiments are heightening attention to named objects or merely heightening infant attention (which may or may not be reflected in the amount of looking time accrued during the familiarisation phase.) The only way to distinguish between a general attention heightening effect of novel labels and a heightening of attention to named objects is to demonstrate that infants are encoding an object–label mapping. And to show this, it is necessary to demonstrate that infants map *distinct* labels to *distinct* objects (Schafer & Plunkett, 1998). None of the experiments reviewed above demonstrate that a specific object–label (or category-label) mapping has been acquired by the infant.

Do words highlight the commonalities between objects? Waxman and Braun (2005) report a novelty preference task with 12-month-old infants familiarised with a sequence of four objects accompanied by the same label, different labels or a no-word condition. An out-of-category novelty preference was only observed in the consistent labelling condition, leading the authors to conclude that ‘‘naming distinct objects with the *same* name highlights the commonalities among them and supports categorization’’ (B66). However, Waxman and Braun did not evaluate infant performance in a no-sound condition. Consequently, it is not possible to dismiss the possibility that their results reflect a tendency for inconsistent labelling to interfere with categorisation, rather than for consistent labelling to highlight commonalities between objects.

Evidence from a series of studies (Booth & Waxman, 2003; Waxman, 1999; Waxman & Booth, 2001, 2003) has led these researchers to argue that ‘‘by 14 months of age, infants may have an emerging expectation linking adjectives specifically to object properties’’ whereas they ‘‘expect count nouns to refer to object categories’’ (Booth & Waxman, 2003, p. 377). Whilst this claim is an interesting and important corollary to the claim that labels highlight commonalities between objects, we do not agree that their studies offer clear evidence to support their position. Although, a detailed discussion of these studies is beyond the scope of the present article, the studies show similar shortcomings to the other studies mentioned previously. We also would point out that in order to show that adjectives highlight common properties and that nouns highlight category membership, it is necessary that infants are tested under identical conditions following distinct conditions of familiarisation and that these lead to *opposite* patterns of responding during the testing phase. None of the studies achieve this outcome.

In summary, we find little evidence for the claim that labels facilitate categorisation during infancy. At best, the evidence suggests that novel words enhance infant attention to the experimental stimuli. However, in the absence of a ‘‘silent’’ control in the overwhelming majority of studies, even this claim lacks solid support. Similarly, the claim that labels highlight the commonalities between objects is found lacking, given the absence of baseline controls and/or appropriate test comparisons. At best, the evidence indicates that labels can *re-activate* existing categories.

More generally, a common methodological limitation permeates all the studies reviewed above: Infants are only ever familiarised with a single category of objects and a single label. Under these experimental conditions, an object category is never independently motivated by the label and so it is impossible to decide whether any observed effect is purely attentional or is motivated by an association between object and label. In order to determine whether a label–object association can highlight the commonalities between objects, it is necessary to familiarise the infant with at least *two* labels and then demonstrate that the structure of the labelling contingencies has had an impact on infant categorisation. Alternatively, it is possible to demonstrate that *one* label can have an impact on categorisation if infants are exposed to stimuli where they would normally form *two categories* but in the presence of just a single label a *new* category is formed. Under these circumstances, the experimenter may be justified in claiming that label–object associations have an impact on infant categorisation. This is the evidence we provide in the following series of experiments.

1.2. Overview of experiments

A convincing demonstration that labels can have an impact on category formation requires that new categories of objects have been learned by the infant and that the structure of the acquired categories was shaped by labels contingent with those categories. We conducted a series of novelty preference experiments using the stimuli depicted in Fig. 1. They were taken from Younger's (1985) report of 10-month-old infants' categorisation of cartoon drawings of animals. A cartoon drawing assumed any of five values along each of four dimensions corresponding to neck and leg length, tail size and ear orientation. Hence, drawing 1155 in Fig. 1 has short legs, long neck, small tail and spread ears, whereas drawing 5511 has long legs, short neck, large tail, and upright ears. Infants were assigned to a Broad or Narrow familiarisation condition. In the Broad condition, values on one dimension could combine with the full range of values on other dimensions, e.g., a long necked drawing was just as likely to have short legs as long legs. However, in the Narrow condition values on one dimension were predictive of values on other dimensions, e.g., spread ears predicted small tails and *vice versa*. The Narrow condition invited infants to group the drawings into two categories: One category consisting of drawings with combinations of low values on each dimension and the other category consisting of combinations of high values. The Broad condition drawings were indicative of just a single category of objects whose average (or prototype) was represented by the draw-

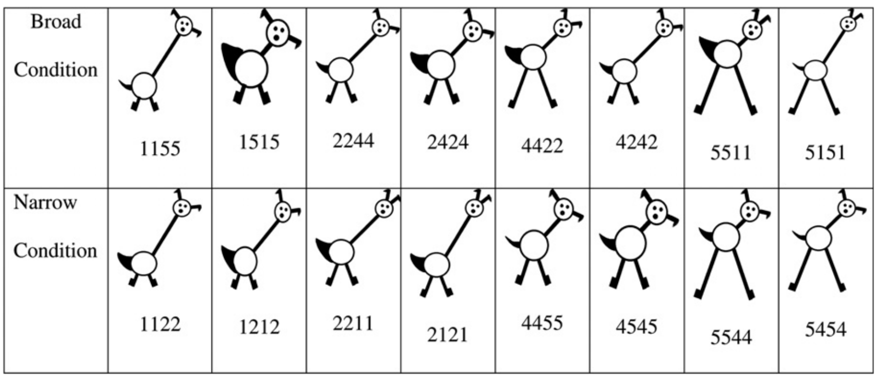

Fig. 1. Two sets of line drawings used in the familiarisation phase of experiments. The Broad condition drawings were used in Experiment 1. The Narrow condition drawings were used in Experiments 2–5.

ing 3333, an exemplar not displayed during the familiarisation phase of the experiments.

We, like Younger (1985), tested category formation using a novelty preference task. The basic assumption, verified by more than four decades of research (Fantz, 1964; Roder et al., 2000; Wetherford & Cohen, 1973), is that infants will look longer at novel than at familiar stimuli. We reasoned that infants familiarised with Broad condition drawings (our Experiment 1) would recognise drawing 3333 (see Fig. 2) as familiar since it represented the average of all the drawings presented whereas drawings that combined just low or high values, such as 1111 or 5555, would appear relatively novel. In contrast, infants familiarised with Narrow condition drawings (our Experiment 2) would find the drawings 1111 or 5555 representative of the two narrow categories and 3333 novel because it belonged to neither of the categories formed during familiarisation.

Thus, infants tested after familiarisation with Broad condition stimuli should look longer at 1111 or 5555 than at 3333, and infants in the Narrow condition should do the reverse. These were precisely the results reported by Younger (1985) and our first step was a successful replication of her original findings using computer generated graphics and eye-tracking methods that offer a high degree of precision in the measurement of eye fixations.

We then ran three additional experiments (our Experiments 3, 4 and 5). In each case, we familiarised infants with the Narrow (two category) condition stimuli, but in each experiment the infants experienced a different use of labels during familiarisation.

Experiment 3: Two labels, one for each category, with all members of each category sharing the same label.
Experiment 4: Two labels with each object pseudo-randomly assigned one of the labels (thus, two tokens of each label were presented with each category).
Experiment 5: One label used throughout, with every drawing given the same label.

Category formation was tested in exactly the same way for all experiments, i.e., following familiarisation, infants were shown an average stimulus (3333) versus a category representative stimulus (1111 or 5555). No labels were presented during this test phase. We predicted that infants in Experiment 3 should replicate Younger's original Narrow condition results since the two familiarisation labels correlated with

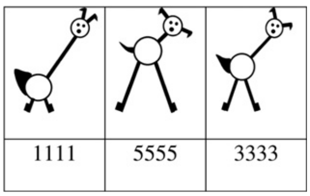

Fig. 2. The test stimuli used in all experiments.

category membership, i.e., they should form two categories and therefore, exhibit a novelty preference for the average stimulus. If labels have an impact on category formation in 10-month-olds, then in Experiment 4, the random assignment of labels should disrupt any category formation, i.e., infants should not exhibit any systematic novelty preference. However, in Experiment 5, the single, consistently used label should invite infants to form a single category, as in Younger's original Broad condition, i.e., infants should exhibit a preference for the 1111 (or 5555) stimulus.

2. Methods

Infants were recruited through the maternity ward at the John Radcliffe Hospital in Oxford. British English was the only language spoken at home and none of the infants had any known history of visual or auditory impairment. We tested up to 33, 10-month-old infants (mean age 10 months 9 days, range 9 months 21 days to 10 months 12 days) in each experiment. After drop-outs due to infant fussiness (see Table 1), the data from 24 infants were available for each experiment.

Infants sat on their parent's lap facing a back projection screen, at a distance of approximately 100 cm. Whilst infants watched the visual stimuli on the back projection screen, parents listened to instructions played over headphones against a background of white noise and were asked to close their eyes so that they could not bias their infant's behaviour. The visual and auditory stimuli were delivered from digitally recorded files stored on a presentation computer, using custom written software that permitted precise timing in the presentation of the auditory and visual files. The auditory stimuli were delivered via a loudspeaker located centrally above the projection screen.

Two hidden cameras, placed immediately above the displayed images, were used to create a digital video recording of infant eye fixations on the visual stimuli. The digital video recording was synchronised with a log file created by the presentation computer to specify the precise onset and offset of familiarisation and test trials. Eye fixations were scored offline from the digital video recording on a frame-by-frame basis. There was no online coding of infant looking. Each digital frame in the recorded trials was coded as left, right or elsewhere depending on whether the infant was fixating the left location, the right location or neither, respectively. The coder was blind to the status of the trial (familiarisation vs. testing), the left/right location of the visual stimuli and the identity of the auditory stimuli (if any were used). The trials were coded in the same sequence as was presented to the infant. A custom written analysis programme then scanned the tagged digital video recording to provide

Table 1
Participant details for Experiments 1–5

Experiment	1	2	3	4	5
Participants	24 (11M + 13F)	24 (12M + 12F)	24 (12M + 12F)	24 (13M + 11F)	24 (12M + 12F)
Drop-out (rate)	9 (27.2%)	8 (25%)	4 (14.2%)	5 (17.2%)	5 (17.2%)

an accurate (to 40 ms) estimate of looking time to the displayed images. Inter- and intra-judge scoring reliability for the frame-by-frame analysis of the digital videos was conducted for 15% of the infants and yielded correlations >0.97.

Each experiment consisted of a familiarisation phase, immediately followed by a test phase. In the familiarization phase, infants were shown eight line drawings identical to one of the sets depicted in Fig. 1. During familiarisation, each line drawing was displayed, one at a time, to the left or the right (chosen randomly by the computer) of the centre of the back projection display, immediately below one of the hidden cameras, for a fixed period of 10 s. The line drawing covered an area of 28×34 cm and its inside edge was 9.5 cm from the centre of the display screen. Order of presentation of the line drawings was randomised by the presentation computer at run time. The inter-trial interval (ITI) was infant controlled. If the infant was fixating the display screen when a trial ended, the experimenter immediately triggered the start of the next trial. This yielded a minimum ITI of about 1 s. If the infant's attention drifted between trials, the experimenter triggered a visual attention getter to centre the infant's eye gaze. On average, the familiarisation phase of the experiment lasted about 90 s.

The test phase of each experiment occurred immediately after the last familiarisation trial. Infants were shown pairs of test objects, side by side, consisting of individual pictures they had not seen before but similar to the familiarisation set (see Fig. 2). Nineteen centimeter separated the pictures. One of the novel objects was always 3333, the drawing with the average value across all dimensions. The other novel object was either 1111 or 5555 (randomly selected across participants), i.e., one of the drawings depicting the extreme values on each dimension. In Experiments 1 and 2, the test pair was displayed for 10 s (replicating Younger's original timing) and then repeated with the left–right positioning of the pair reversed. The ITI during the test phase of the experiment was infant controlled, in the same manner as the familiarisation phase. Maintenance of the same timing parameters used in the original Younger study permitted a clear evaluation of the impact of the different methods of stimulus presentation (a Fagan apparatus vs. computer controlled, digitally recorded stimuli) on the robustness of Younger's original findings.

In Experiments 3–5, test pairs were displayed for 6 s each. We chose to decrease the length of the test trials in Experiments 3–5 because of the relatively high drop-out rate observed in the first two experiments (see Table 1). Furthermore, we aimed to maximise the proportion of total time infants spent fixating the two pictures during testing: Time spent looking away from the pictures does not contribute to the measure of novelty preference and may be detrimental to the infant's engagement in the task. Note that the drop-out rate was considerably lower in Experiments 3–5 (see Table 1), though this may also be attributed to the presence of the auditory stimuli during the familiarisation phase of the experiments.

The auditory stimuli heard during familiarisation in Experiments 3–5 consisted of an initial carrier phrase "Look!" presented 2 s after the onset of the visual stimulus which initially appeared in silence. The label ("dax" or "rif") assigned to the visual stimulus was then played twice, once 5 s and then 8 s after the onset of the visual stimulus. Each label was pre-recorded in an enthusiastic female voice using infant directed speech. As stated previously, no auditory stimuli were presented during the test phase in any experiment. Likewise, there were no auditory stimuli played during the familiarisation phases of Experiments 1 and 2, thereby replicating Younger's original design.

3. Results and discussion

3.1. Familiarisation phase

We analysed the time infants spent attending to objects during familiarisation by calculating the average amount of time they spent fixating each of eight familiarisation stimuli. In addition, we calculated the average time infants fixated the first three familiarisation stimuli (Block 1) and the final three familiarisation stimuli (Block 2). A comparison of average looking times in Blocks 1 and 2 enabled us to determine whether infants showed any signs of habituation during the familiarisation phase of the experiments. Table 2 presents the overall average looking times and the average looking times for Blocks 1 and 2 in all five experiments. A comparison of the overall looking times across experiments suggests that infants spent more time looking at the visual stimuli during familiarisation in the presence of an auditory stimulus (Experiments 3–5) than when the cartoon figures were presented in silence (Experiments 1 and 2). Furthermore, a comparison of average looking times in Blocks 1 and 2 indicates a reduction in looking time during familiarisation for all experiments, suggesting that some habituation has occurred. A 5×2 mixed model ANOVA with the factors Experiment (five levels) and Block (two levels) revealed a main effect of Experiment ($F(4,115) = 24.886$, $p < .001$; $\eta^2 = .464$) and Block ($F(1,115) = 31.891$, $p < .001$; $\eta^2 = .217$). The interaction term Experiment $\times$ Block was not significant ($F(4,115) = .878$, $p = .479$ (n.s.); $\eta^2 = .030$). Post hoc comparisons revealed that infants spent longer time looking at the familiarisation drawings in Experiment 1 than Experiment 2 ($F(1,115) = 9.65$, $p = .002$; $\eta^2 = .077$), and that they looked longer during familiarisation in Experiments 3–5 than both Experiments 1 and 2 (all ps $< .001$). These findings confirm that infants spent more time looking at the visual stimuli in the presence of an auditory stimulus than when it was absent, and that they showed some habituation during the familiarisation phase of the experiments.

Table 2
Mean looking time (s) in the familiarisation trials

	N	Block 1	Block 2	Grand mean/*SD*
Experiment 1 (Broad condition)	24	5.810 (1.558)	4.639 (1.297)	5.225 (1.236)
Experiment 2 (Narrow condition)	24	4.337 (1.788)	3.605 (1.672)	3.971 (1.578)
Experiment 3 (Narrow condition)	24	7.298 (1.735)	6.450 (2.052)	6.874 (1.691)
Experiment 4 (Narrow condition)	24	7.438 (1.313)	7.043 (1.365)	7.240 (1.078)
Experiment 5 (Narrow condition)	24	7.386 (1.228)	6.721 (1.602)	7.054 (1.452)

Standard deviation in parentheses.

The results for the three auditory experiments (3–5) in Table 1 all indicate comparable overall looking times and familiarisation effects, thus suggesting that the variation in the number of different labels heard during familiarisation (two in Experiments 3 and 4 and one in Experiment 5) had little or no impact on looking time. This was confirmed in 3×2 mixed model ANOVA with factors Experiment (three levels) and Block (two levels) which revealed a main effect of Block ($F(1, 69) = 12.861$, $p < .001$, $\eta^2 = .157$). There was no main effect of Experiment nor significant interaction between Experiment and Block.

3.2. Test phase

We analysed the time infants fixated each picture in the test phase of each experiment and calculated the proportion of time infants fixated the average stimulus (3333) compared to the total time fixating both stimuli (3333/(3333 + 1111|5555)). These results are presented in Table 3. In Experiment 1 (no-sound, Broad condition), infants spent more time looking at the objects with extreme values (1111 or 5555) than the object with average values (3333), while in Experiment 2 (no-sound, Narrow condition) the opposite pattern of results was found. Both of these looking preferences were significantly different from chance (0.50), and replicated the pattern of results from Younger's original study. We can conclude that the infants in these two experiments were treating the stimuli in the same way as the infants in Younger's original study: Exposure to the Broad set of objects in the absence of any auditory stimulus resulted in the formation of a single category, whereas exposure to the Narrow set yielded two distinct categories.

Experiment 3, where the two labels correlated with visual category structure, produced the same pattern of results as Experiment 2. Infants showed a significant novelty preference for object 3333 over 1111 or 5555 ($t(23) = 2.55$, $p < .02$, $d = 0.47$) indicating that they had formed two categories. Note that a direct statistical comparison of the novelty preferences in Experiments 2 and 3 is not warranted since the duration of the test phases are different (10 and 6 s, respectively). However, the average proportional preferences for object 3333 are very similar in the two experiments, suggesting that the labels did not have a facilitating effect in the categorisation

Table 3
Looking times and statistical analyses for Experiments 1–5

Experiment	1	2	3	4	5
Looking time 3333:1111/5555 (s)	2.76:3.35	3.43:2.72	2.30:1.90	2.14:2.04	1.88:2.14
Percentage of time on average 3333 (*SD*)	44.17 (9.55)	56.14 (11.81)	55.09 (9.77)	50.28 (12.19)	45.75 (7.72)
t tests (2-tailed)	$t(23) = -2.99$	$t(23) = 2.55$	$t(23) = 2.55$	$t(23) = 0.11$	$t(23) = -2.70$
p value	.007	.018	.018	.912	.013
Effect size (Cohen's *d*)	0.61	0.52	0.47	0.02	0.55

Note that test trials last 10 s in Experiments 1 and 2 and 6 s in Experiments 3–5.

process. Were infants simply ignoring the two auditory labels or monitoring their correlation with the set of visual objects? Experiment 4 reveals what is happening in Experiment 3. In Experiment 4, where the two labels are randomly assigned to individual cartoon figures, infants showed no novelty preference whatsoever: Their responding was not systematically different from chance ($t(23) = 0.11$, $p = .912$, $d = 0.02$). Since the only difference between Experiments 3 and 4 is the pattern of correlation between the two labels and the eight visual stimuli, Experiment 4 shows that infants must have noticed the correlation between use of the two labels and the category structure even in Experiment 3.

Experiment 4 shows that labels can disrupt the formation of categories otherwise supported by the structure inherent in the visual domain. The results of Experiment 5 show that labels can override and impose a new structure on the visual domain. In Experiment 5, where a single label is used for all familiarisation items, infants produced a significant preference for objects 1111 and 5555 over 3333 ($t(23) = -2.70$, $p = .01$, $d = 0.55$), indicating that they had formed just a single category. This result mimics the findings of Experiment 1 where infants were trained with visual stimuli from the Broad condition, but without any labelling. In contrast, the infants in Experiment 5 were trained on Narrow condition stimuli, and yet they responded as if they have formed a single category. Moreover, the novelty preference observed in Experiment 5 was the opposite to that observed in Experiment 2 where exactly the same visual stimuli were presented in the absence of labels during familiarisation. Although, a direct statistical comparison of Experiments 2 and 5 is not warranted due the difference in the duration of the test phases, the contrast in novelty preferences highlights the causal role of the label in the process of category formation in Experiment 5.

4. General discussion

It was argued earlier that a necessary condition to demonstrate that labels have a facilitative impact on infant categorisation is through comparison with infant categorisation of the same stimuli in the absence of auditory input. Furthermore, it was argued that a necessary condition to demonstrate that labels play a causal role in category formation, rather than merely highlight attention to stimuli, is to show that infants pay attention to label–object associations during the process of category formation and that the correlational structure of these associations influence the type of categories constructed by the infant. A convincing way to achieve this is to require that the infant identify at least two label–object associations.

Experiments 1 and 2 demonstrated that 10-month-old infants can readily construct novel categories of objects in the absence of auditory stimulation. Evidence for category formation was provided in a novelty preference test. Infants showed contrasting patterns of novelty preference depending on the pattern of correlation between the features that defined the familiarisation objects. When features presented broad patterns of correlation, infants formed a single category, evidenced by a novelty preference for the extreme combinations of object features (1111 or

5555). When features presented narrow patterns of correlation, infants formed two categories, evidenced by a novelty preference for the combination of object features that fell between the categories (3333). These findings replicate those of Younger (1985), reported over 20 years ago, using a very different testing apparatus (computer controlled stimulus presentation with offline scoring vs. a Fagan apparatus with online scoring), indicating that these infant categorisation findings are robust and can be observed using quite different methods.

Infants' attention to the familiarisation stimuli was enhanced by the presence of an auditory label. On average, infants spent 3 s longer attending to the visual stimuli, out of a possible 10 s, when accompanied by auditory stimuli than when presented in silence. Note that the shorter looking times in Experiments 1 and 2 were not due to higher levels of habituation. Infants showed similar amounts of attenuation in looking time across all five experiments (there was not a significant interaction between Block and Experiment). The impact of the auditory stimuli on overall looking times during familiarisation supports the claim that novel labels heighten attention to objects, or at least to pictures of objects (Waxman, 1999).

The results of Experiments 3–5 show that novel labels not only heighten attention to pictures of objects but that the structure of the label–object contingencies during familiarisation has an impact on infant category formation. In Experiment 3, infants were familiarised with a set of eight visual stimuli, which clustered into two categories, together with two labels which correlated with category membership. At test, infants demonstrated a reliable novelty preference for object 3333, thereby demonstrating that they had constructed two categories during familiarisation. This result enables us to conclude that the presentation of the two novel labels did not interfere with the category formation process: At test, infants performed in a manner similar to Experiment 2 where they provided evidence of having formed two categories in the absence of any auditory stimuli during familiarisation. Note that we cannot conclude that the novel labels in Experiment 3 facilitated category formation, despite the fact that novel labels heightened infant attention to the objects during familiarisation. A direct statistical comparison of the levels of novelty preference in Experiments 2 and 3 was not warranted because the duration of the test phases differed in the two experiments. However, the proportional measures of preference for the average stimulus were very similar in the two experiments (56% vs. 55%) suggesting that the presence of the labels in Experiment 3 did not enhance the novelty preference.

One interpretation of the combined results of Experiments 2 and 3 is that the labels in Experiment 3 have no causal role other than to heighten infants' general attention. The results of Experiment 4 rule out this interpretation and demonstrate that infants paid attention to the correlational structure of label–object associations during familiarisation and that infant sensitivity to these correlations was influential in the category formation process. The outcome of Experiment 4 was that infants showed no evidence of a novelty preference which indicates that they failed to form two categories (or indeed a single category centred on the prototype). It is important to emphasise that exactly the same eight objects and two labels were used during familiarisation in Experiment 4 as Experiment 3. The only thing that changed was the pattern of label–object associations in the two experiments. Yet, the novelty preferences expressed by the infants were systematically different. Therefore, we can conclude that the random assignment of labels to objects in Experiment 4 was responsible for infants' failure to form any object categories. In other words, infants must have monitored the correlation between the two labels and the visual category structure in *both* experiments. Taken together, Experiments 3 and 4 demonstrate that infants can readily compute the cross-modal correlations between label–object associations and object–feature correlations in the process of category formation. If the cross-modal correlations align with intra-modal visual feature correlations, then category formation is supported. However, if the cross-modal correlations differ from the intra-modal correlations, then category formation is disrupted.

The results of Experiment 5 show that labels can play a constructive role in infant category formation. In this experiment, infants heard a single label together with a set of objects which perceptually clustered into two categories. Indeed, in the absence of any auditory stimulation, infants formed two categories during the course of familiarisation (Experiment 2). However, the consistent labelling of the same visual stimuli in Experiment 5 with just one auditory stimulus ('dax' or 'rif') overrode the visual category structure to produce a reversal of the novelty preference, indicating that infants had formed just a single category. This finding is not only consistent with the claim that labels highlight the commonalities between objects and facilitate the categorisation process but it indicates that labels can override *dissimilarities* between objects so that they are treated as being perceptually more similar.

We are able to make these claims because the addition of the label in Experiment 5 results in a *different* category structure from the one that is formed in the absence of a label. The label in Experiment 5 is not merely highlighting attention to the objects as the default response is to segregate the objects into two categories. Nor is the label merely interfering with the processing of the visual stimuli since this would lead to an absence of a novelty preference (as in Experiment 4). Experiment 5, together with Experiment 2, shows that infants monitored the label–object associations during familiarisation and that these associations played a causal role in category formation.

5. Conclusion

The current series of experiments provides unequivocal support for the view that labels impact the process of categorisation in young infants even before they begin to produce their first words, to the extent that labels can override the perceptual dissimilarities between objects and lead infants to treat them as more similar to each other. The experiments also demonstrate that young infants can simultaneously compute the correlational structure of object features in the visual domain *at the same time* as they compute the relationship of that correlational structure to novel features (words) in the auditory domain. This cross-modal, computational capacity should prove a powerful tool for the young infant to exploit in deriving the meaning of words.

Our findings would appear to contrast with another body of research which maintains that auditory information (including words) can *overshadow* the processing of visual information by young infants (e.g., Robinson & Sloutsky, 2004). According to this approach, overshadowing occurs when the auditory stimuli are relatively novel, so that infants will dishabituate to an auditory change in an habituated, compound auditory–visual stimulus but not to a visual change. The current study did not attempt to establish full habituation during the familiarisation phase of the experiments, so direct comparison with the Robinson and Sloutsky (2004) findings is inappropriate. However, the current findings, together with those of Waxman and colleagues, indicate that *novel* auditory stimuli need not overshadow the processing of visual structure. Otherwise, we would have failed to observe novelty preferences between the visual stimuli which were tested in silence. Indeed, our results demonstrate that young infants can compute the correlational structure of a series of visual stimuli at the same time as they are exposed to novel words. Far from overshadowing the visual stimuli, the novel auditory stimuli in these experiments were directly associated with the visual stimuli in a manner that impacted the process of object categorisation itself.

The capacity of labels to impact upon the categorisation process in these experiments may seem sufficiently potent to suggest that they act as the primary motivators of category structure. Indeed, the experiments show that infants may pay close attention to the correlational structure inherent in the label–objects associations to which they are exposed, and that labels can encourage infants to unite visual stimuli into a single category that would otherwise be treated as two categories in the absence of any labels. However, the series of experiments described in this paper does not address the cases where labels might motivate infants to *divide* objects into discrete categories when there is no evidence for a categorical structure in the visual stimuli or when the categorical structure of the visual stimuli is opposite to that suggested by the labels. For example, consider the case of the Broad condition stimuli (see Fig. 1) that promote the formation of one category in the absence of any labels (as in Experiment 1). Would it be possible to familiarise infants with these stimuli together with two distinct labels so that infants form two categories amongst the visual objects? In other words, can labels both unite *and* divide visual categories? Answers to questions such as these would enable us to establish what constraints may operate in determining the impact that labels can have on infant categorisation processes.

We suggest that the methodology exploited in the set of experiments described in this paper offers a principled approach to discovering the role that language might play in shaping infant categorisation. To reiterate, convincing demonstrations that labels can impact the processes of infant categorisation require experimental outcomes showing that the categorisation process is contingent upon the structure of the labelling events. Invariably, this will require that at least two visual categories and/or two labels be tested in order to demonstrate that the category formation process is motivated by the label–object associations themselves and is not merely a result of heightened attentional processing completely unrelated to the labelling contingencies.

Whatever the outcome of further investigation, these experiments offer a firm empirical foundation for the claim that labels can impact the process of infant categorisation. The implications of this claim are profound. Insofar as we have demonstrated that labels can encourage infants to unite discrete visual categories into a single category (Experiment 5), we may conclude that caregiver labelling practises will help infants appreciate the similarities between visually dissimilar events and thereby highlight the categories that are important in the infant's native culture. The type of visual events that might be susceptible to the impact of the caregiver's labelling practises are diverse, ranging from social visual stimuli such as facial and emotional expressions (helping an infant identify what counts as, say, an angry or a happy face) to pure perceptual categories such as colour (helping an infant or young child discover what hues count as "green" in her native culture).

We have had little to say about the range of labelling events that might impact the process of infant categorisation. Our auditory stimuli were restricted to phonotactically legal, monosyllabic CVC labels. Would a non-linguistic, auditory stimulus such as a bell or a buzzer, or even additional visual stimuli such as different coloured backgrounds, have a similar impact on category formation? Studies indicating that linguistic stimuli do not have a privileged status in early word learning (e.g. Woodward & Hoyne, 1999) would suggest that a wider range of stimuli might be effective in manipulating infant category structure. Notwithstanding the privileged status (or lack thereof) of words in early infancy, their ubiquitous application in the labelling practices of adults will ensure that words will play a central role in the way that infants form categories during their first year of life.

References

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-olds? *Journal of Experimental Child Psychology 64*, 3–26.

Behl-Chadha, G. (1996). Basic-level and superordinate-like categorical representations in early infancy. *Cognition, 60*, 105–141.

Booth, A. E., & Waxman, S. (2002). Object names and object functions serve as cues to categories for infants. *Developmental Psychology, 38*(6), 948–957.

Booth, A. E., & Waxman, S. R. (2003). Mapping words to the world in infancy: Infant's expectations for count nouns and adjectives. *Journal of Cognition and Development, 4*(3), 357–381.

Eimas, P. D., & Quinn, P. C. (1994). Studies in the formation of perceptually based basic-level categories in young infants. *Child Development, 65*, 903–917.

Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science, 146*, 668–670.

Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not tones) facilitate object categorization: Evidence from 6- and 12-month-olds. *Cognition, 105*, 218–228.

Fulkerson, A. L., & Haaf, R. A. (2003). The influence of labels, non-labeling sounds, and source of auditory input on 9- and 15-month-olds' object categorization. *Infancy, 4*(3), 349–369.

Quinn, P. C., & Eimas, P. D. (1996). Perceptual organization and categorization in young infants. In C. Rovee-Collier & L. P. Lipsitt (Eds.). *Advances in infancy research* (Vol. 10, pp. 1–36). Norwood, NJ: Ablex.

Quinn, P. C., & Johnson, M. H. (2000). Global-before-basic object categorization in connectionist networks and 2-month-old infants. *Infancy, 1*, 31–46.

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its changed in the course of development. *Child Development, 75*, 1387–1401.

Roder, B. J., Bushnell, E. W., & Sasseville, A. M. (2000). Infant's preferences for familiarity and novelty during the course of visual processing. *Infancy, 1*, 491–507.

Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development, 69*(2), 309–320.

Waxman, S. R. (1999). Specifying the scope of 13-month-olds' expectations for novel words. *Cognition, 70*, B35–B50.

Waxman, S. R., & Booth, A. E. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology, 43*, 217–242.

Waxman, S. R., & Booth, A. E. (2003). The origins and evolution of links between word learning and conceptual organization: New evidence from 11-month-olds. *Developmental Science, 6*, 130–137.

Waxman, S. R., & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition, 95*, B59–B68.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29*, 257–302.

Wetherford, M., & Cohen, L. B. (1973). Developmental changes in infant visual preferences for novelty and familiarity. *Child Development, 44*, 416–424.

Woodward, A. L., & Hoyne, K. L. (1999). Infant's learning about words and sounds in relation to objects. *Child Development, 70*(1), 65–77.

Younger, B. A. (1985). The segregation of items into categories by 10-month-old infants. *Child Development, 56*, 1574–1583.