Journal of Experimental Child Psychology 123 (2014) 147–154

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Brief Report

Three-year-olds obey the sample size principle of induction: The influence of evidence presentation and sample size disparity on young children's generalizations

Chris A. Lawson

Department of Educational Psychology, University of Wisconsin–Milwaukee, Milwaukee, WI 53201, USA

A R T I C L E I N F O

Article history:
Received 2 August 2013
Revised 5 December 2013
Available online 17 January 2014

Keywords:
Sample size
Induction
Generalization
Sequential presentation
Simultaneous presentation
Ratio difference

A B S T R A C T

Three experiments with 81 3-year-olds (M = 3.62 years) examined the conditions that enable young children to use the sample size principle (SSP) of induction—the inductive rule that facilitates generalizations from large rather than small samples of evidence. In Experiment 1, children exhibited the SSP when exemplars were presented sequentially but not when exemplars were presented simultaneously. Results from Experiment 3 suggest that the advantage of sequential presentation is not due to the additional time to process the available input from the two samples but instead may be linked to better memory for specific individuals in the large sample. In addition, findings from Experiments 1 and 2 suggest that adherence to the SSP is mediated by the disparity between presented samples. Overall, these results reveal that the SSP appears early in development and is guided by basic cognitive processes triggered during the acquisition of input.

Introduction

Young children are remarkably gifted inductive reasoners insofar as they are able to generalize a single piece of evidence about an individual (e.g., "this rabbit has omat inside") to an entire class of individuals (e.g., "rabbits have omat inside") (for a review, see Gelman, 2003). However, early induction appears to be limited in important ways. For example, evidence indicates that several principles of induction emerge sometime after 7 years of age (Gutheil & Gelman, 1997; Li, Cao, Li, Li, & Deak,

E-mail address: lawson2@uwm.edu

0022-0965/$ - see front matter
http://dx.doi.org/10.1016/j.jecp.2013.12.004

2009; Lopez, Gelman, Gutheil, & Smith, 1992). The current studies examined the emergence of the *sample size principle* (SSP), which dictates that large samples of evidence provide better evidence from which to generalize than do small samples (Osherson, Smith, Wilkie, Lopez, & Shafir, 1990). Specifically, three experiments explored the possibility that the SSP is available early in development, but only under conditions that facilitate the detection of differences, rather than similarities, between samples of evidence.

The SSP is typically studied in sample comparison tasks in which participants are asked to generalize to a novel animal either a property associated with a large sample of animals (e.g., 5 mammals have "property *y*") or a different property associated with a small sample of animals (e.g., 3 mammals have "property *x*"). Under such conditions, by around 8 years of age children exhibit a preference to generalize the property associated with the large sample (e.g., property *y*), whereas younger children show no preference to generalize from either sample (Lopez et al., 1992). In contrast, children younger than 6 years exhibit more robust generalizations for large samples than for small samples when sample size is manipulated between participants rather than within participants (Jacobs & Potenza, 2001; Lawson & Fisher, 2011). These mixed findings suggest that adherence to the SSP might be evidence dependent rather than age dependent.

The current studies examined two factors that can affect whether children incorporate sample size information into their inductive decisions. The first is the method of evidence presentation. At least one study showing early adherence to the SSP involved sequential presentation of items (e.g.. Lawson & Fisher, 2011), whereas studies showing later development of the SSP involved simultaneous presentation of items (Gutheil & Gelman, 1997; Li et al., 2009; Lopez et al., 1992). Simultaneous presentation supports alignment of shared features (Gentner & Namy, 1999), whereas sequential presentation supports detection of differences between the presented items (Lappin & Bell, 1972; Quinn & Bhatt, 2010). Because comparison of differences, rather than similarities, between items is a necessary condition for the SSP, sequential presentation might better facilitate the SSP than simultaneous presentation. In addition, sequential presentation of exemplars might confer an information processing advantage for large samples of input. For example, sequential presentation supports consolidation of multiple items in visible short-term memory, whereas simultaneous presentation does not (Liu & Becker, 2013). Thus, sequential presentation of evidence facilitates another necessary condition of SSP—detection of multiple items in the larger sample. A primary goal of the current studies was to examine the potential impact of these different presentation formats on children's adherence to the SSP.

The preference to generalize from large samples may also be influenced by the size disparity between presented samples. Humans exhibit an early ability to detect differences between two samples by relying on the size disparity between samples (Brannon, 2002; Halberda & Feigenson, 2008; Lipton & Spelke, 2003). Evidence of later emergence of the SSP comes from studies that presented smaller and less disparate samples (e.g., 5 vs. 3 in Lopez et al., 1992) than studies showing early emergence of SSP (e.g., 30 vs. 1 in Jacobs & Narloch, 2001). Thus, a second goal of the current studies was to examine whether differences in use of the SSP can be linked to children's sensitivity to the size disparity between samples.

Finally, because there were no developmental predictions, these experiments used a single age group (3-year-olds).

Experiment 1

In this experiment participants were told about a property attributed to a small sample of animals (e.g., a cat with unti inside) and a different property attributed to a large sample of animals (e.g., several cats with omat inside) and then were asked to generalize one of the properties to a target (e.g., a cat). Presentation format (sequential and simultaneous) was manipulated between participants to test the prediction that sequential presentation would elicit a higher rate of generalizations from the large sample than would simultaneous presentation.

In addition, the number of items in each sample was manipulated such that the size of the large sample varied (e.g., 2, 3, 4, and 5 exemplars), whereas the small sample always had 1 exemplar.

The hypothesis was that children would be more likely to generalize from the large sample for the more disparate samples (e.g., 5 vs. 1) than for the less disparate samples (e.g., 2 vs. 1).

Method

Participants

The sample of 30 3-year-olds (M = 3.57 years, SD = 0.27, 16 girls and 14 boys) comprised 12 Caucasians, 8 African Americans, 5 Hispanics/Latinos, 2 Asian Americans, and 3 children identified by their guardians as "other." None of the participants was included in Experiment 2 or 3.

Design, materials, and procedures

There were eight trials, each of which included a large sample, a small sample, and a target. All items were from the same basic-level category (e.g., cats). Four categories (cats, dogs, insects, and fish) were each presented twice, and toy figurines (approximately 4–6 cm in diameter) were used to represent each animal. The number of individuals in the large sample varied, whereas the small sample always had one exemplar, thereby creating the following sample pairs: 2 versus 1 (2v1), 3v1, 4v1, and 5v1. Each sample pair was presented twice. Items were blocked such that each sample pair and category was presented once in each of two blocks.

Each trial included the attribution of a novel property to the evidence samples (see Appendix for a list of properties). In the Sequential condition (n = 15) each animal was presented individually and attributed a novel property (e.g., "This animal has [Property 1]") and then placed in a pile to the left of the participant. For the next sample each individual was attributed a different property than was attributed to individuals in the first sample (e.g., "This animal has [Property 2]") and then placed in a pile to the right of the participant. In the Simultaneous condition (n = 15) individuals were not presented individually. Instead, both the large and small samples were presented at the same time, one to the left of the participant and the other to the right of the participant. The participant was then told about one sample, "this/these animal/s [Property 1]," during which time the experimenter pointed to the pile. Next, the experimenter pointed to the other sample and described "this/these animal/s [Property 2]." In both conditions the samples were presented approximately 40 cm apart from each other and approximately 30 cm from the participant. The presentation of items was counterbalanced so that each sample (large or small) was presented first (and to the left of the participant) a nearly equal number of times.

Both evidence samples were left visible while the participant was asked to make generalizations to an individual from the same basic-level as the individuals in the evidence samples (e.g., a different cat). The participant was presented with a figurine and asked, "Do you think this animal [Property 1] like this/these animal/s or [Property 2] like this/these animal/s?" Each target was presented approximately 15 cm in front of the child and equidistant to both evidence samples. After the child responded, the next trial was presented.

Results and discussion

Initial analysis of responses in Experiments 1 to 3 confirmed that there was not a preference to generalize based on the side to which a sample was presented or the order in which a sample was presented, all $Fs < 1.20$, *ns*.

To test the main hypotheses, the proportion of generalizations of the property attributed to the large sample were submitted to a mixed analysis of variance (ANOVA) with evidence presentation (Sequential or Simultaneous) as a between-participants variable and sample pair (2v1, 3v1, 4v1, or 5v1) as a within-participants variable. Both effects were significant, as was the evidence presentation by sample pair interaction, $F(3, 84) = 3.76$, $p = .01$, $\eta_p^2 = .12$. As indicated in Fig. 1, there was a higher rate of generalizations from the large sample in the Sequential condition than in the Simultaneous condition for the 5v1 and 4v1 pairs, both $Fs > 10.61$, $p < .002$. The Sequential condition yielded an increase in generalizations from large samples as a function of sample pair disparity, $F(3, 42) = 9.03$, $p < .001$, $\eta_p^2 = .39$. In the Simultaneous condition, there were no significant differences between any ratio pairs (M range = .47–.58), $Fs < 1$, *ns*.

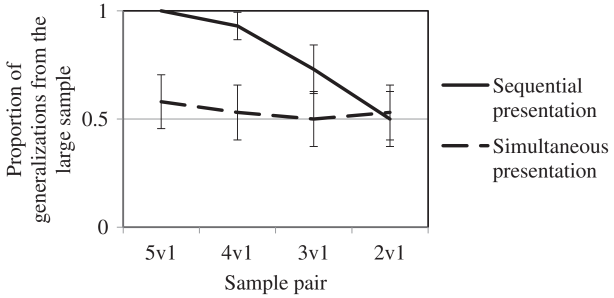

Fig. 1. Proportion of generalizations from the large sample in the Simultaneous and Sequential conditions for each of the four sample pairs (2v1, 3v1, 4v1, or 5v1) in Experiment 1. Bars indicate 1 standard error from the mean.

The final analyses compared responses to chance ($M = .50$), with the assumption that scores above chance indicate a preference to generalize from the large sample, whereas scores below chance indicate a preference to generalize from the small sample. Participants never exhibited a preference to generalize from the small sample. With exception to the 2v1 pair, the Sequential condition elicited a consistent preference to generalize from the large sample, all ts > 2.82, ps < .02, ds > 1.50. Responses in the Simultaneous condition were no different from chance, all ts < 1.20, *ns*.

These results support the two main hypotheses. First, children exhibited the SSP when evidence was presented sequentially but not when it was presented simultaneously. Second, the sample size effect appeared for sample pairs with large disparities (e.g., 5v1) but not for those with small disparities (e.g., 2v1). However, because the small sample always had one individual, it is not clear whether the sample pair effect was due to the relative size of the large sample or the size disparity between samples. Furthermore, the source of the sequential presentation effect remains unclear. Was the effect due to differences in presentation time or to different cognitive processes triggered by the presentation format? The following experiments explored these issues.

Experiment 2

This experiment replicated the Sequential condition in Experiment 1 with the exception that participants were presented with the following sample pairs: 5v1, 5v2, 5v3, and 5v4. If the SSP is mediated by the disparity between samples, there should be an incline in the rate of generalizations from the large sample as samples become more disparate. In contrast, if participants are drawn to the relative size of the large sample, there should be no difference in the rates of generalization for each of the sample pairs because the large sample always included 5 exemplars.

Method

Participants

The sample of 21 3-year-olds ($M = 3.54$ years, $SD = 0.43$ years, 13 girls and 8 boys) comprised 11 Caucasians, 6 African Americans, and 4 Hispanics/Latinos. None of these participants was included in Experiment 1 or 3.

Design, materials, and procedure

The design was identical to the Sequential condition in Experiment 1 with the exception that children were presented with two items from each of the following sample pairs: 5v1, 5v2, 5v3, and 5v4.

Results and discussion

The proportion of generalizations from the large sample was submitted to a sample disparity (5v1, 5v2, 5v3, or 5v4) ANOVA, which yielded a significant effect, $F(3, 57) = 4.83$, $p = .005$, $\eta_p^2 = .20$.

Consistent with the sample disparity prediction, generalizations from the large sample were higher for the 5v1 ($M = .80$) and 5v2 ($M = .68$) pairs than for the 5v3 ($M = .53$) and 5v4 ($M = .43$) pairs, all $Fs > 2.30$, $ps < .04$. In addition, the rate of generalizations from the large sample in the 5v1 and 5v2 pairs was significantly greater than would be expected by chance, both $ts > 2.61$, $ps < .02$, $ds > 1.19$, whereas generalizations for the less disparate sample pairs were not different from chance, $ts < 1$, *ns*.

These results replicate and extend the findings from Experiment 1; sequential presentation elicited the SSP in 3-year-olds, and this effect was mediated by the disparity between the presented samples rather than the relative size of the large sample.

Experiment 3

One explanation for the results in Experiments 1 and 2 is that the presentation effect was due to the longer exposure to items in the Sequential condition compared with the Simultaneous condition. For example, the additional time required to present items may have supported encoding of, or attention to, all of the input. If this was the case, the effect should be absent when exposure to items is held constant across conditions. Experiment 3 tested this prediction.

Another interpretation is that different processes triggered by sequential and simultaneous formats have different consequences for induction. As noted earlier, sequential presentation facilitates memory consolidation for multiple items, whereas simultaneous presentation does not. The SSP in the sequential condition may be due to a processing advantage for the larger sample of input such that enhanced memory for more items in this sample than in the small sample makes the large sample the better inductive base from which to generalize. This interpretation suggests that, in addition to facilitating the SSP, the sequential presentation might lead to greater recognition of items from the large sample than does the simultaneous presentation. Experiment 3 explored this possibility.

Method

Participants

The sample of 30 3-year-olds ($M = 3.74$ years, $SD = 0.43$, 17 girls and 13 boys) comprised 13 Caucasians, 10 African Americans, 6 Hispanics/Latinos, and 1 Asian American. None of these participants was involved in Experiment 1 or 2.

Design, materials, and procedure

The design was modeled after Experiment 1 with two modifications. First, presentation time was controlled to ensure that there was equal exposure to items in the Simultaneous and Sequential conditions. To ensure that children sustained attention to the items in the Simultaneous condition, the experimenter repeated the prompts three times; approximately every 4 s, the experimenter mentioned while pointing to the appropriate samples, "these/this animal/s [Property 1] and this/these animal/s [Property 2]." The second modification was the addition of a memory task that consisted of two sample items (e.g., two cats presented in the large sample), two close foils (e.g., two cats not presented in the samples), and two far foils (e.g., a pig and a cow). The memory task was administered after all generalization items were presented. For each item, the participant was presented with six figurines and asked, "Can you show me which animals you saw in the game we just played and which you did not see?"

Finally, all generalization items were presented in 5v1 and 5v2 pairs because these sample pairs yielded consistent sample size effects in Experiment 2.

Results and discussion

An initial analysis confirmed that presentation time in the Sequential condition ($M = 17.17$ s) was not significantly different from presentation time in the Simultaneous condition ($M = 16.78$ s), $t < 1$, *ns*. Despite similarities in presentation time, children were more willing to generalize from the large sample when items were presented sequentially ($M = .73$) than when items were presented simultaneously ($M = .47$), $F(1, 29) = 8.52$, $p = .007$. Supplemental analyses indicated that the rate of generalizations from large samples for both sample pairs was consistently greater than chance in the Sequential condition, both $ts > 3.90$, $ps < .01$, $ds > 2.00$, whereas generalizations for both pairs were not different from chance in the Simultaneous condition, both $ts < .50$, *ns*.

For the memory probe, the main question of interest was whether the presentation effects were linked to better memory for the sample items. Analysis showed that the proportion of times children correctly recognized the sample items (hits) was higher in the Sequential condition ($M = .84$) than in the Simultaneous condition ($M = .59$), $F(1, 29) = 8.69$, $p = .006$, $d = 3.22$. Supplemental analysis found that the proportion of times children recognized the foils (false positives) was no different in the two conditions, $F < 2.00$, *ns*, thereby suggesting that the sequential condition did not simply lead to a greater willingness to judge that an item had been seen.

Consistent with the results from Experiments 1 and 2, the sequential presentation elicited inferences in accordance with the SSP. Moreover, these presentation effects were not due to differences in the duration of exposure to the evidence items. Instead, greater recognition for the sample items in the sequential presentation indicates that use of the SSP in this condition is facilitated, at least in part, by how input is processed when exemplars are presented in serial order.

General discussion

The goal of these studies was to determine whether mixed findings concerning the onset of the SSP of induction reflect genuine age-wise changes in inductive capacity or whether they are due to children's sensitivity to different features of the available evidence. The results from three experiments with 3-year-olds provide strong support for the latter. These young children exhibited the SSP when evidence was presented sequentially rather than simultaneously and when there was sufficient disparity in the size of the presented samples. Thus, these results indicate that this inductive principle emerges early in development, but only under certain evidential conditions.

One interpretation of these results is that sequential presentation, unlike simultaneous presentation, elicits specific processes that support adherence to the SSP. For example, sequential presentation supports better memory for multiple items in the large sample thereby making this the better sample from which to generalize. However, because the samples were always left visible during the generalization task, better memory alone is insufficient to account for the results. For example, recognition of items in the sample could result from greater attention to, or deeper processing of, these items during the presentation phase. Indeed, sequential presentation likely engages several processes that can influence how children evaluate multiple samples of evidence to make an inductive generalization. A second interpretation of the presentation effects considers the different comparison processes; sequential presentation supports the identification of sample differences, whereas simultaneous presentation encourages comparison of similarities between items and, thus, inhibits use of the SSP. Rather than identifying a specific mechanism that supports adherence to the SSP, the results from the current studies suggest that a constellation of processes that jointly support the detection of multiple items within samples and encourage comparison of differences between samples contributes to children's use of sample size as a basis for making inductive generalizations.

Further, the results suggest that adherence to the SSP does not require understanding that large samples provide better coverage of the category represented in the samples (e.g., Osherson et al., 1990). From the category-based perspective children fail to use the SSP because they are unable to represent the category shared by the evidence samples and targets (Gutheil & Gelman, 1997; Lopez et al., 1992). It could be argued that the sequential presentation of items facilitated representation of the category covered by the evidence samples. However, such an interpretation lacks parsimony without appealing to the processes that contributed to the formation of the appropriate inclusion category.

The current work aligns with a large body of work showing that children are flexible inductive reasoners (Hayes & Thompson, 2007; Kalish & Gelman, 1992; Opfer & Bulloch, 2007; Sloutsky & Fisher, 2008). The idea that task dynamics influences children's performance on cognitive tasks is not new (e.g., Trabasso et al., 1978), and the effects of simultaneous and sequential presentation formats have been explored in the developmental literature on word learning (Spencer, Perone, Smith, & Samuelson,

2011) and relational reasoning (Son, Smith, & Goldstone, 2011). However, this is the first study to directly compare the effects of these presentation formats on children's inductive reasoning. The sequential versus simultaneous paradigm has generated a rich set of conclusions about different cognitive and perceptual processes, many of which can be useful for understanding the development of induction. Thus, an important direction for future work is to examine the impact of these presentation methods on other aspects of inductive reasoning.

A second novel finding was that adherence to the SSP was influenced by the size disparity between evidence samples such that children preferred to generalize from larger samples for the more disparate sample pairs. This result extends findings in the number cognition literature (e.g., Brannon, 2002; Lipton & Spelke, 2003) by suggesting that sensitivity to ratio differences between samples can serve as a cue for making inductive generalizations. However, there is disagreement about whether sensitivity to ratio differences is guided by attention to the number of items in the evidence or by other properties such as the spatial extent of the stimuli (e.g., Clearfield & Mix, 1999). Likewise, it remains to be seen whether the SSP reflects a preference for a numerically larger sample or for the sample that is physically more expansive or complex—or whether something else draws attention to the large sample or away from the small sample. The SSP has been operationalized, and assessed, in a variety of ways, such that more stringent measures than reveal limitations in adults' use of the SSP (e.g., Kahneman & Tversky, 1972). Although the current studies reveal the early presence of some form of the SSP, more work is needed to understand the development and scope of this important reasoning principle.

In sum, the results from the current study indicate that adherence to the SSP is mediated by specific features of the evidence that are known to support the detection of differences between exemplars. Thus, rather than representing a late-developing inductive skill, the sample size principle of induction appears early in development under cases that elicit basic cognitive processes available to quite young learners.

Appendix A. Properties used in Experiments 1 to 3

Property 1	Property 2
Eats unti food	Eats grigon food
Has omat inside	Has unti inside
Has a 3-chambered heart	Has a 2-chambered heart
Has plaxium blood	Has drotium blood
Drinks lapy water	Drinks juto water
Has auxin inside	Has umblat inside
Has dinot eyes	Has gunit eyes
Has noben cells	Has polik cells

References

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition, 83*, 223–240.
Clearfield, M., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science, 10*, 408–411.
Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York: Oxford University Press.
Gentner, G., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14*, 487–513.
Gutheil, G., & Gelman, S. A. (1997). Children's use of sample size and diversity information within basic-level categories. *Journal of Experimental Child Psychology, 64*, 159–174.
Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology, 44*, 1457–1465.
Hayes, B. K., & Thompson, S. P. (2007). Causal relations and feature similarity in children's inductive reasoning. *Journal of Experimental Psychology: General, 136*, 470–484.
Jacobs, J. E., & Narloch, R. H. (2001). Children's use of sample size and variability to make social inferences. *Applied Developmental Psychology, 22*, 311–331.

Jacobs, J. E., & Potenza, M. (2001). The use of judgment heuristics to make social and object decisions: A developmental perspective. *Child Development, 62*, 166–178.
Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*, 430–454.
Kalish, C. W., & Gelman, S. A. (1992). On wooden pillows: Multiple classification and children's category-based inductions. *Child Development, 63*, 1536–1557.
Lappin, J. S., & Bell, H. H. (1972). Perceptual differentiation of sequential visual patterns. *Perception & Psychophysics, 12*, 129–134.
Lawson, C. A., & Fisher, A. V. (2011). It's in the sample: The effect of sample size on the breadth of inductive generalization. *Journal of Experimental Child Psychology, 110*, 499–519.
Li, F., Cao, B., Li, Y., Li, H., & Deak, D. (2009). The law of large numbers in children's diversity-based reasoning. *Thinking & Reasoning, 15*, 388–404.
Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science, 14*, 396–401.
Liu, T., & Becker, M. W. (2013). Serial consolidation of orientation information into visual short-term memory. *Psychological Science, 24*, 1044–1050.
Lopez, A., Gelman, S. A., Gutheil, G., & Smith, E. E. (1992). The development of category-based induction. *Child Development, 63*, 1070–1090.
Opfer, J. E., & Bulloch, M. J. (2007). Causal relations drive young children's induction, naming, and categorization. *Cognition, 105*, 206–217.
Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97*, 185–200.
Quinn, P. C., & Bhatt, R. S. (2010). Learning perceptual organization in infancy: The effects of simultaneous versus sequential variability experience. *Perception, 39*, 795–806.
Sloutsky, V. M., & Fisher, A. V. (2008). Attentional learning and flexible induction: How mundane mechanisms give rise to smart behaviors. *Child Development, 79*, 639–651.
Son, J. Y., Smith, L. B., & Goldstone, R. L. (2011). Connecting instances to promote children's relational reasoning. *Journal of Experimental Child Psychology, 108*, 260–277.
Spencer, J. P., Perone, S., Smith, L. B., & Samuelson, L. K. (2011). Learning words in space and time: Probing the mechanisms behind the suspicious-coincidence effect. *Psychological Science, 22*, 1049–1057.
Trabasso, T., Isen, A. M., Dolecki, P., McLanahan, A. G., Riley, C. A., & Tucker, T. (1978). How do children solve class inclusion problems? In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp 151–180). Hillsdale, NJ: Lawrence Erlbaum.