Cognition 150 (2016) 243–251

Contents lists available at ScienceDirect

Cognition

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

Discussion

No evidence of learning in non-symbolic numerical tasks – A comment on Park and Brannon (2014)

Marcus Lindskog [*,1], Anders Winman [1]

Uppsala University, Uppsala, Sweden

ARTICLE INFO

Article history:
Received 1 October 2014
Revised 27 December 2015
Accepted 7 January 2016

Keywords:
Numerical cognition
Approximate Number System
Cognitive training

ABSTRACT

Two recent studies – one of which was published in this journal – claimed to have found that learning on a non-symbolic arithmetic task improved performance on a symbolic arithmetic task (Park & Brannon, 2013, 2014). This finding has potentially far-reaching implications, because it would constitute evidence for a causal link between the Approximate Number System (ANS) and symbolic-math ability. Here, we argue that, due to the methodology used in both studies, the interpretation of data in terms of an improvement in ANS performance is problematic. We provide arguments and simulations showing that the trends in the data are similar to what one would expect for a non-learning observer. We discuss the implications for the original interpretation in terms of causality between non-symbolic and symbolic arithmetic performance.

1. Introduction

The Approximate Number System (ANS) is thought to be a primitive cognitive system that supports the representation of non-symbolic magnitudes (e.g., Feigenson, Dehaene, & Spelke, 2004). It has been documented in human adults (e.g., Halberda, Ly, Wilmer, Naiman, & Germine, 2012), infants (e.g., Feigenson et al., 2004), and non-human animals (e.g., Brannon, Wusthoff, Gallistel, & Gibbon, 2001). Several studies have indicated that having a more precise ANS is related to better arithmetic ability (e.g., Halberda, Mazzocco, & Feigenson, 2008; Inglis, Attridge, Batchelor, & Gilmore, 2011; Libertus, Feigenson, & Halberda, 2011). This finding has attracted a lot of interest and suggested a causal functional link.

In two recent studies Park and Brannon (henceforth P&B, 2013, 2014) propose that the ANS is causally related to symbolic-math ability. The claim is supported by experimental demonstrations of transfer of learning from a non-symbolic arithmetic, to a symbolic arithmetic task in terms of a math test. P&B (2013, p. 2015) suggested that the results of their study "...show that improvement in an ANS-based, nonsymbolic, approximate-arithmetic training task over multiple sessions transfers to selective improvements in symbolic-math ability." In P&B (2014) the scope was widened by use of several tasks that measured various cognitive components that might be responsible for a causal effect. This strategy aimed at "improving distinct cognitive components" (p. 189) in order to later "compare the transfer effects in exact symbolic arithmetic performance across these training conditions" (p. 189).

These results potentially have very important implications for our understanding of human numerical cognition. For example, Hyde, Khanum, and Spelke (2014, p. 93) argued that the findings of P&B (2013) "...provide the strongest evidence to date of a causal and specialized relationship between the ANS and symbolic mathematics." From an applied perspective, implications are overwhelming. As suggested by P&B (2014, p. 199) the results could mean that approximate arithmetic training could be used in society to "benefit young children who have yet to master the meaning of exact number or numerical symbols".

In both studies, the main conclusions critically depend on the finding that performance improvements on an ANS-based task transferred to a symbolic math task. The logic behind this would be that if X is causally related to Y, an improvement due to training of ability X should induce an alteration of ability Y. More specifically, if it were possible to show that an improvement in ANS-performance by training is accompanied by a subsequent improvement in symbolic arithmetic performance it would suggest a direction of causality from ANS to symbolic math ability. Here, however, we argue that it is not possible to interpret the data provided in the two studies by P&B as showing any improvement in non-symbolic arithmetic at all. We provide simulations suggesting

DOI of original article: http://dx.doi.org/10.1016/j.cognition.2016.02.012

* Corresponding author at: Department of Psychology, Uppsala University, P.O. Box 1225, SE-751 42 Uppsala, Sweden.

E-mail address: marcus.lindskog@psyk.uu.se (M. Lindskog).

[1] M. Lindskog and A. Winman contributed equally to this work.

http://dx.doi.org/10.1016/j.cognition.2016.01.005
0010-0277/

that the trends in data that P&B interpreted as evidence for learning are the trends that one would expect to find for non-learning observers.

2. Park and Brannon's adaptive training method

In the approximate-arithmetic task used by P&B, participants see two arrays of dots moving behind an occluder in sequence (addition task) or one array of dots moving behind an occluder followed by another array of dots appearing from behind the occluder (subtraction task). Participants evaluate the result of the operation (addition or subtraction) implied by the movements of the dot arrays. The responses from participants are elicited in two ways. On comparison trials, participants are presented with a new array of dots and decide if the result of the previously observed operation is more ore less numerous than the new array. On match trials, participants are presented with two new arrays and decide which of the two match the numerosity of the result of the operation. The difficulty of the task is determined by the ratio between the correct answer and the alternative response option.

P&B used an adaptive method, similar to those used in psychophysics (e.g., the "up-down method") to estimate individuals' psychophysical discrimination threshold on various tasks (see e.g., Treutwein, 1995), to train their participants. In their implementation of the method, task difficulty is adjusted after every 20 trials according to how well the participant performed: if performance on the last 20 trials was above 85% correct, the difficulty is increased; if it was below 70% correct, the difficulty is decreased. P&B found that task difficulty stabilized after several sessions, at a level that was considerably harder than the initial one (Fig. 1A, filled dots). Their conclusion from this observation was that performance had improved, indicating learning.[2]

The seemingly very rapid change of stimulus difficulty as a function of training found in all adaptive tasks in both studies (P&B, 2013, 2014) is at first glance striking. From the first to the second training session participants in the different conditions seemingly master visual short-term memory tasks with a higher span, a symbol ordering task at higher speed and an approximate arithmetic task with stimuli much harder to discriminate. Most impressive are maybe the effects found on Approximate Number Comparison, which seem to suggest that performance of participants dramatically improved within 25 min of training with a reduction of weber fractions by two thirds. This finding is surprising considering that other studies (e.g., Lindskog, Winman, & Juslin, 2013) have tried without success to obtain learning by training in very similar tasks (see also DeWind & Brannon, 2012). P&B (2013) suggested that a possible explanation of this discrepancy may lie in the regulation procedure (the adaptive algorithm) "which kept the task challenging", thereby "inducing active engagement" (p. 2017) within participants. While we agree that the regulation procedure embedded in the adaptive algorithm is important per se in understanding the findings, we propose a different explanation than actively engaged participants.

3. "Improvement" without learning

P&B interpreted the increase in difficulty level during training (Fig. 1A, filled dots) as evidence that subjects had gotten better at the task. However, with the adaptive method used by P&B, the direction of convergence (harder/easier) critically depends on the starting value chosen by the experimenter. This intuition is demonstrated in Fig. 1A that illustrates the results of a simulation where participants with equal performance but different starting values take on the task used by P&B. The figure, shows that a relatively easy starting value (filled squares) necessary will lead to convergence on harder stimuli (lower values on the y-axis) whereas a hard starting value (open squares) will bring about convergence on easier stimuli (high values on the y-axis) (see full details about simulations below).

The starting level chosen by P&B was easier than what has been found to be readily mastered by 6-month-old human infants (ratio 1:2) (e.g., Starr, Libertus, & Brannon, 2013; Xu, 2003; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005). Hence, it is not surprising that they found that the adaptive method converged to more difficult stimulus levels over time – this is what one would expect, even for observers who do not learn. Had they used a relatively difficult starting level, they would possibly have found – with the same observers – a decrease of difficulty over time. Therefore, the direction of convergence cannot be used to determine whether participants got better at the task.

To establish that learning has taken place, one could instead conduct pre- and post-tests in combination with a proper control group. P&B (2014) did actually obtain pre- and post-tests for three of the measures for which they claim improvement took place with the adaptive tests; Approximate Number Precision, Visuospatial short-term memory, and Numeral order Judgments. Albeit not identical, these tests were very similar to the adaptive tests. Thus, one would expect near transfer effects on these tests if the observed pattern of performance on the adaptive tests were due to learning. No such effects in terms of increased accuracy were found on the pre-post comparisons (an effect was found in terms of faster reaction times on the numerical symbol ordering task).[3] In spite of this finding, the authors interpreted the changes in stimulus difficulty on the adaptive tests in terms of learning.

4. A simulation of a non-learning observer

Another way to establish whether learning has taken place is to compare the human data with data from simulations of a non-learning observer. If the trends in the human data are very different from those predicted for non-learning observers, then the human data may be interpreted as evidence for learning. On the other hand, if simulations of a non-learning observer closely mimic the human data, then it is questionable to conclude that the human data contain evidence for learning.

To investigate what data would look like for a non-learning observer, we performed simulations of Experiments 1 and 2 in P&B (2013) and Experiment 1 in P&B (2014), which we will refer to as E1-2013, E2-2013, and E1-2014, respectively. In these simulations, we used the same procedures as P&B, except that simulating an ideal observer, instead of collecting it from a human observer, generated the response on each trial.

Following previous work (e.g., Barth et al., 2006; Dehaene, 2001; Dehaene & Changeux, 1993) we assume that numerosity estimates are internally represented on a logarithmic scale with constant Gaussian noise. Hence, if we denote the numerosity of a stimulus by N, then the simulated internal representation of this numerosity, n, is drawn from a Gaussian distribution with a mean equal to $\log(N)$ and a standard deviation σ. We further assume that the observers use the optimal decision rule to make their choices.

[2] The present paper focuses on the interpretation of performance in an approximate arithmetic task. However, P&B (2013, 2014) make claims about improvement in performance for tasks involving "numerical ordering", "approximate number comparison", "short-term memory" and "symbol ordering". The main objections presented below of such an interpretation likewise fully apply to all these other tasks.

[3] In order to try to demonstrate near transfer effects, P&B performed post hoc contrasts pooling the approximate arithmetic and the non-symbolic numerical comparison groups. Those analyses approached statistical significance, but did not reach the conventional alpha level of .05.

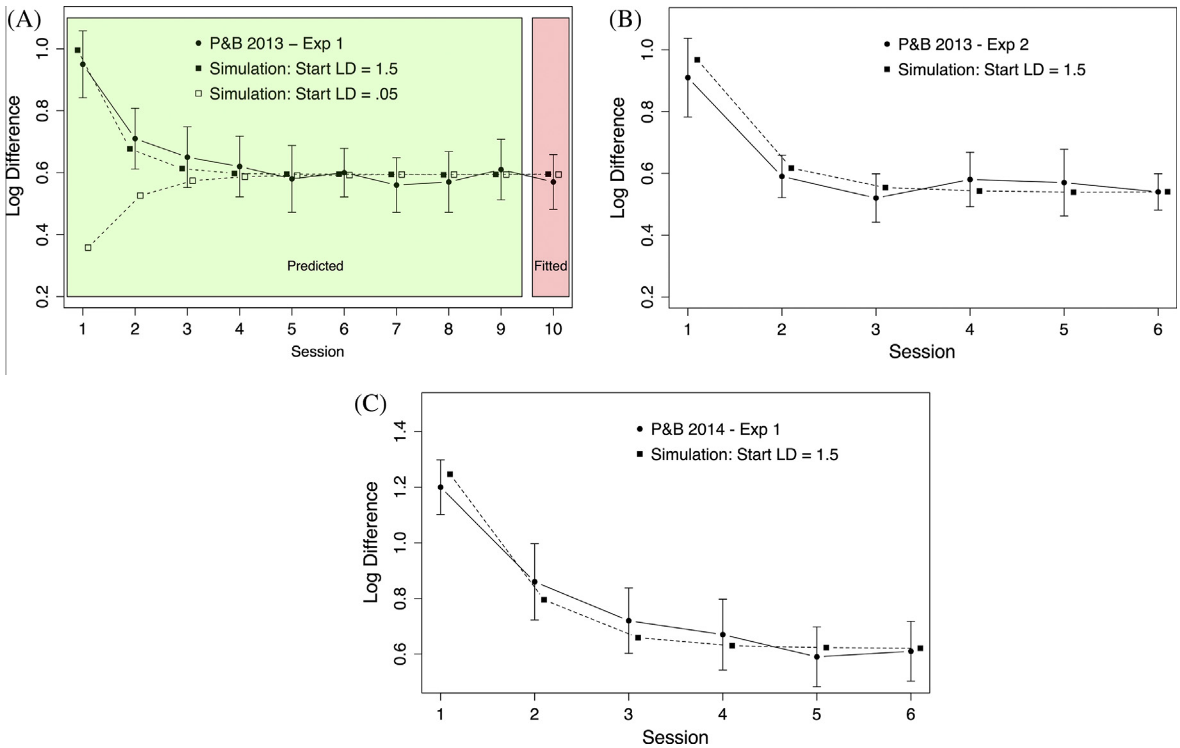

Fig. 1. Results from a simulation with a non-learning, ideal observer. The ideal observer accurately predicts the empirical trend in each of the three experiments. (A) Experiment 1 from P&B (2013). (B) Experiment 2 from P&B (2013). (C) Experiment 1 from P&B (2014). Vertical bars denote 95%-confidence intervals. We chose to report 95%-CI rather than the standard errors reported by P&B to ease the interpretation.

In the comparison task, participants are briefly presented with two numerosities, N_1 and N_2, and decide if their difference, N_1–N_2, is larger or smaller than a probe, N_3.[4] In this task, an ideal observer responds "larger" when $(n_1 - n_2) > n_3$ and smaller otherwise. In the match task, participants are briefly presented with two numerosities, N_1 and N_2, and respond which of two probes, N_3 or N_4, has equal numerosity to the difference between N_1 and N_2. Here the ideal observer responds "N_3" if $|n_3 - (n_2 - n_1)| < |n_4 - (n_2 - n_1)|$ and "N_4" otherwise. The level of observation noise, σ, is the only parameter of the model. Learning in the model is determined by the time-dependence of this parameter: a learning observer is one whose value of σ decreases over time, while a non-learning observer is one whose value of σ is fixed over time.

To examine if P&B's human data are consistent with predictions of a non-learning observer, we applied the following method. First, we determined for each of the three experiments which value of σ provides a good match to the very last empirical data point, where observers had reached a stable difficulty level that gives 77.5% correct performance – the asymptotic level of performance in P&B – (see Fig. 1A, red shaded area for an illustration). We found that suitable values of σ are .22, .20, and .23 for experiments E1-2013, E2-2013, and E1-2014, respectively. Next, we fixed σ to these values and simulated P&B's adaptive procedure to obtain model predictions (see Fig. 1, green[5] shaded area for an illustration) for the entire experiments (to obtain accurate predictions, we averaged the results over 10,000 runs).

P&B defined stimulus difficulty as the numerical distance, on a $\log_2$-scale, (log-difference, *LD*) between the result of the operation and the probe(s). As in the original studies our simulated participants started their training at a LD of 1.5 and completed 100 (60, 60) blocks of 20 trials each. Half of the trials were simulated as comparison trials and half as match trials, in accordance with P&B's procedure. For each block we simulated participants' performance by randomly drawing values from the distributions corresponding to the internal representation of numerosity in the respective tasks and applying the decision rule of the ideal observer. The level of difficulty for a block was determined by performance on the previous block as described by P&B. In accordance with P&B's procedure we adjusted the LDs for the comparison and match tasks separately.

Fig. 1 illustrates the results of the simulations together with the results reported by P&B.[6] As can be seen in the figure, the averaged model predictions closely match the averaged human data in every experiment. Notice that although the precision parameter in the simulations was estimated from the last data point in the respective experiments in P&B, the simulations also closely mimics the data with respect to the starting point, the rate of decay, and the asymptotic performance. This means that the trends in the data reported by P&B are those that one would expect for a non-learning observer and, thus, that there is no need at all to assume learning in order to explain the pattern of data reported by P&B.

5. Discussion

We have argued that it is not possible to draw the conclusion of an improvement in performance from data collected with an adaptive test procedure, and illustrated that the results of P&B look very much as they would with non-learning participants. It is important to note that it is possible that participants nevertheless did

[4] In P&B participants also perform an addition task, which make the same predictions with respect to performance as the subtraction task formulated here and which is easily derived from the subtraction case.

[5] For interpretation of color in Fig. 1, the reader is referred to the web version of this article.

[6] Results from P&B were extracted with WebPlotDigitizer (Rohatgi, 2014).

improve and that a real improvement in quantity manipulation could occur without any accompanying improvement in the performance measures at all. However, as our demonstration shows, there is nothing in the data reported by P&B that suggests that an improvement took place. The strong causality interpretation made by P&B (2014) in terms of a change through an improvement in "nonverbal numerical quantity manipulation" (p. 188) may nevertheless still be correct. It is of course also possible that an improvement may have occurred on an underlying mechanism that is not tapped by the task performance measure per se, such as for example storing arrays in working memory.

What are the implications for the causality interpretation when learning has not been demonstrated? The possibility that the results have causes that do not depend on a lasting improvement by training of the ANS (be it acuity, quantity manipulation or any other aspect of this system) must be considered. The question is whether or not the transfer effect to symbolic math occurred due to learning caused by training or due to some other more superficial transfer effect independent of learning. It is possible that the mere engaging in non-symbolic arithmetic in absence of any kind of improvement can enhance symbolic arithmetic performance. This is suggested by the results of Hyde et al. (2014) who found effects on children's symbolic math performance after a mere 60 practice trials of approximate arithmetic in the absence of non-symbolic task improvement. It seems implausible that such a negligible amount of practice can have made a change in the effectiveness with which children manipulate quantities at all. The authors speculated that this effect was due to common cognitive mechanisms engaged in these tasks. It may for example be the case that approximate numeric processes "warm up" relevant neurological areas, and that this "priming" is transferred to the symbolic arithmetic test. This explanation may seem less plausible considering that the post-test math tests in P&B (2013, 2014) took place on a different session than the last training session 1-2 days later. However, priming effects may be very persistent. For example, Mitchell (2006) found priming effects after 17 years for a material presented only three times at the original occasion.

A second possibility is that the approximate arithmetic condition promotes a non-specific response speed increase. It is not far-fetched to imagine that for example participants in the probably highly cognitively demanding approximate mental arithmetic task, in order to be able to cope adopt a strategy of responding rapidly after performing the mental arithmetic out of necessity, before the stimuli is displaced in short-term memory. Such an urgency in responding may then be transferred to post-training tasks, and show up in the arithmetic task that depends heavily on fast responding. Effects of non-symbolic arithmetic training were specific to the arithmetic test. However, guessing was punished on the vocabulary tasks and the other accuracy tests consisted of a non-speeded fixed set of items. This fact could have masked response speed effects on those tasks. Consistent with this hypothesis, in analyses of the math task result, P&B (2014) indeed found an increment in the number of math problems attempted, but not for the proportion of these problems solved correctly. Something that may speak against this hypothesis is that the approximate arithmetic group did not exhibit the shortest response times on the numeral order judgment task at posttest. However, there are several important and relevant differences between the numeral order judgment test and the math test. For example, the numeral order judgment test is not timed, but self-paced, with no incentive for participants to respond rapidly. All participants receive a fixed amount of trials irrespective of how long they take to complete the task. In the math test on the other hand, participants finish as many problems as they can within a given time limit, and this determines performance. With this task, faster participants will perform better by just moving from one problem to the next swiftly, something which is not the case for the former task. An item-to-item speed up would thus not show up as an improvement in the numeral order judgment test, which merely measured reaction times after stimulus presentation. Conversely, whereas it might seem plausible that the numerical symbol ordering condition might a priori be especially susceptible to induce a sense of urgency in participants, at the same time this task does not involve discrete items, but a continuous stream of stimuli. In this sense, even if the task is speeded, it might mot induce a sense of urgency because it involves passive reacting to stimuli and unlike the math test no discrete item-to-item progression.

Apparently participants in the different conditions in P&B (2014) received different amounts of rest after a training session depending on the length of the training itself.[7] In the approximate arithmetic condition stimulus presentation takes more than four times as long as it does in the approximate number comparison task. It is quite possible that subtle experimental confounds like this induce different degrees of sense of urgency over conditions so that for example those who received longer breaks felt less sense of urgency than those with briefer periods of rest.

This general speed increase hypothesis[8] may quite straightforwardly be empirically falsified by subjecting participants to some other timed tasks in which performance heavily rests on the speed at which the questions are answered. Future research will have to address the nature of these effects, and should test for general speed effects. If the demonstrated effects in the end are found to be an effect due to common cognitive mechanisms they are undoubtedly interesting from a theoretical perspective. A response speed up strategy is theoretically less interesting. However, both of these alternative explanations to Park and Brannon's hypothesis suggest that effects are likely to be transient and situation-specific. Thus, it is unfortunately quite possible that these findings may convey little or no possibility for applied "interventions for math educators".

Acknowledgments

This research was supported by a grant to Anders Winman from the Swedish Research Council. We are indebted to Mats Lind and Ronald van den Berg for reading and commenting on an earlier version of the manuscript.

References

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition, 98*, 199–222.

Brannon, E. M., Wusthoff, C. J., Gallistel, C. R., & Gibbon, J. (2001). Numerical subtraction in the pigeon: Evidence for a linear subjective number scale. *Psychological Science, 12*, 238–243.

Dehaene, S. (2001). Subtracting pigeons: Logarithmic or linear? *Psychological Science, 12*, 244–246. http://dx.doi.org/10.1111/1467-9280.00343.

Dehaene, S., & Changeux, J.-P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience, 5*, 390–407.

DeWind, N. K., & Brannon, E. M. (2012). Malleability of the approximate number system: Effects of feedback and training. *Frontiers in Human Neuroscience, 6*, 1–10.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*, 307–314.

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive internet-based sample. *Proceedings of the National Academy of Sciences, 109*, 11116–11120.

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*, 665–668.

[7] This information does not appear in the paper, but was conveyed by personal communication with Jonkoo Park.

[8] Note that the term "general speed increase" used here does not imply that participants in any sense have improved in general processing capabilities. It should rather be taken to mean an increased alertness due to experienced time-pressure either due to cognitive demands of the task or external factors in the experimental setting.

Hyde, D. C., Khanum, S., & Spelke, E. S. (2014). Brief non-symbolic, approximate number practice enhances subsequent exact symbolic arithmetic in children. *Cognition, 131*, 92–107.
Inglis, M., Attridge, N., Batchelor, S., & Gilmore, C. (2011). Non-verbal number acuity correlates with symbolic mathematics achievement: But only in children. *Psychonomic Bulletin & Review, 18*, 1222–1229.
Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science, 14*, 1292–1300.
Lindskog, M., Winman, A., & Juslin, P. (2013). Are there rapid feedback effects on approximate number system acuity? *Frontiers in Human Neuroscience, 7*, 270.
Mitchell, D. B. (2006). Nonconscious priming after 17 years: Invulnerable implicit memory? *Psychological Science, 17*, 925–929.
Park, J., & Brannon, E. M. (2013). Training the approximate number system improves math proficiency. *Psychological Science, 24*, 2013–2019.
Park, J., & Brannon, E. M. (2014). Improving arithmetic performance with number sense training: An investigation of underlying mechanism. *Cognition, 133*, 188–200.
Rohatgi, A. (2014). Version 3.3 of webplotDigitizer. *ZENDO*. http://dx.doi.org/10.5281/zendo.10532.
Starr, A. B., Libertus, M. E., & Brannon, E. M. (2013). Infants show ratio-dependent number discrimination regardless of set size. *Infancy, 18*, 927–941.
Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research, 35*, 2503–2522.
Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition, 89*(1), B15–B25.
Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*(1), B1–B11.
Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science, 8*(1), 88–101.

How to interpret cognitive training studies: A reply to Lindskog & Winman

Joonkoo Park [a,b,*], Elizabeth M. Brannon [c]

[a] Department of Psychological and Brain Sciences, University of Massachusetts, United States
[b] Commonwealth Honors College, University of Massachusetts, United States
[c] Department of Psychology, University of Pennsylvania, United States

ARTICLE INFO

Article history:
Received 29 January 2016
Accepted 14 February 2016

Keywords:
Cognitive training
Training effect
Transfer effect
Math intervention

ABSTRACT

In our previous studies, we demonstrated that repeated training on an approximate arithmetic task selectively improves symbolic arithmetic performance (Park & Brannon, 2013, 2014). We proposed that mental manipulation of quantity is the common cognitive component between approximate arithmetic and symbolic arithmetic, driving the causal relationship between the two. In a commentary to our work, Lindskog and Winman argue that there is no evidence of performance improvement during approximate arithmetic training and that this challenges the proposed causal relationship between approximate arithmetic and symbolic arithmetic. Here, we argue that causality in cognitive training experiments is interpreted from the selectivity of transfer effects and does not hinge upon improved performance in the training task. This is because changes in the unobservable cognitive elements underlying the transfer effect may not be observable from performance measures in the training task. We also question the validity of Lindskog and Winman's simulation approach for testing for a training effect, given that simulations require a valid and sufficient model of a decision process, which is often difficult to achieve. Finally we provide an empirical approach to testing the training effects in adaptive training. Our analysis reveals new evidence that approximate arithmetic performance improved over the course of training in Park and Brannon (2014). We maintain that our data supports the conclusion that approximate arithmetic training leads to improvement in symbolic arithmetic driven by the common cognitive component of mental quantity manipulation.

1. Introduction

Imagine yourself in a six-day weight-training program. On the first day, you start squatting 80 lb. Then, you increase the weight adaptively on a daily basis until the last day when you squat 150 lb. Prior to this weight-training program, you could lift up to 200 lb; therefore, technically your *weight-lifting* performance did not improve. Nevertheless, after the six days of squatting, you find that you are able to sprint faster than you previously could! Whether your squatting performance improved or not has little to do with demonstrating the causal relationship between squatting and sprinting and its translational significance (Chelly et al., 2009; McBride et al., 2009). The essence of that causal relationship is *not* between squatting and sprinting but between *strengthening leg muscles* and sprinting.

Lindskog and Winman's (2016) commentary on our previous paper (Park & Brannon, 2014) claim that there is no evidence of *performance improvement*[1] in our non-symbolic approximate

[1] Note that Lindskog and Winman actually argue that there is "no evidence of *learning*." However, they conflate performance improvement in the observable measures with possible changes in unobserved cognitive elements due to training (see Section 2). We suggest to reserve the term learning for the latter, and that it is more accurate to use the term *performance improvement* in this case.

DOI of original article: http://dx.doi.org/10.1016/j.cognition.2016.01.005
* Corresponding author at: Department of Psychological and Brain Sciences, University of Massachusetts, 135 Hicks Way/Tobin Hall, Amherst, MA 01003, United States.
E-mail address: joonkoo@umass.edu (J. Park).

http://dx.doi.org/10.1016/j.cognition.2016.02.012
0010-0277/