Does Variability in an Implicit Learning Task Help Infants Learn Faster? A Registered Report

Emma K. Ward & Sabine Hunnius

Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen,
The Netherlands

Author note: This manuscript was In Principle Accepted as a Stage 1 Registered Report by
Developmental Science on 14 January, 2020.

DOES VARIABILITY HELP INFANTS LEARN?

Research Highlights

- In their everyday lives, infants do not learn from clearly identifiable cues presented in isolation, but from a constant stream of noisy sensory input.
- We will test whether young infants can learn a pattern faster when abstracting away from moderate noise contained in a signal.
- This knowledge will allow us to ultimately develop better theoretical models of how perceptual inference develops and ecologically-valid models of how infants learn.

DOES VARIABILITY HELP INFANTS LEARN?

Abstract

Infants are constantly inundated with sensory input, which they must somehow interpret in order to understand the world. They must do this without explicit cues about which parts of the input are going to be useful and which parts can be ignored. We know that infants do implicitly learn complex skills, which shows that they manage to separate informative signal from uninformative noise.

In laboratory-based learning tasks, when the signal to be learnt is accompanied by other less reliable signals, this environmental variability seems to cue learners towards the invariant, learnable features and allows them to perform better when tested (e.g. Gómez, 2002; Tummeltshammer & Kirkham, 2013). Here, we build on these previous studies by asking whether variability in the exact realisation of events – noise within the signal to be learnt – is also influencing learning during the stimulus presentation. We will test whether this type of variability is, over and above a helpful cue, a crucial and fundamental ingredient for learning that has so far been absent from most experimental paradigms.

Using multi-level modelling to analyse infants' implicit learning during a novel saccadic serial reaction time task, we aim to show how infants learn a sequence from noisy instantiations of events. If infants learn sequences with added noise faster than those without, we will claim that noise acts on expectations online to shape perception as it happens.

Key words: Perceptual inference, implicit learning, eye-tracking, variability, signal and noise, multi-level modelling

The data that support the findings of this study will be openly available on the Open Science Framework after data collection is complete.

Does Variability in an Implicit Learning Task Help Infants Learn Faster? A Registered Report

Our senses receive a continuous stream of information that we must learn to interpret in order to understand the world. By using our previous experiences and expectations, we use perceptual inference to separate the informative signal from uninformative noise, and discern the cause of our sensory input; to form a percept (Aggelopoulos, 2015; Clark, 2013). Only once we have formed a percept in this way can we reason about and learn from the information we are receiving. We define noise here as any input that is not strictly necessary to improve learning, either because it is irrelevant at the moment or because it is unlearnable random variation, for example due to our imperfect senses. Although there are no explicit cues to teach infants which aspects of sensory information are signal and which are noise, they seem to separate the two well from early on, as evidenced by the speed with which they learn complex skills incidentally, such as those seen in language acquisition (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997) and action perception (Reddy, Markova, & Wallot, 2013).

For example, when infants are about to be picked up, they witness someone reach two outstretched hands towards them. Although they see many different sizes of hands, angles of arms and paths of movement, they learn to expect to feel the tactile stimulus of hands on their torso, followed by the proprioceptive change of their body being moved. The exact details of the tactile and proprioceptive input are also different each time the infant is picked up, but by three months old, infants have learnt to disregard these small variations and act upon their knowledge by preparing to be picked up; tensing their neck muscles to prevent their heads from falling backwards (Reddy et al., 2013).

Here, we propose that abstracting away from such variations is not only necessary in order to learn, but that the variations actively support learning through shaping perceptual inference. Variation as it so far has been studied is limited in its scope and has been operationalised more akin to distractors than to changes allowing for rule abstraction. We believe that the role of noise in the signal is equally important to understand. Noise is the specific information surrounding a signal which is along the stimulus dimension of interest but is not crucial to interpret the percept. In contrast, variability is seen as a general property of an environment. A highly-variable environment is one that contains more unique events, however a highly-variable environment may or may not contain more noisy signals,

depending on whether the variations are along the dimension of interest and whether they meaningfully contribute to the task the observer is trying to accomplish.

Variability during training has helped adults and children to generalise better than their peers who have received less variable training sets (Douvis, 2005; Gómez, 2002; Gonzalez & Madhavan, 2011; Wulf & Schmidt, 1997; Yao, Desola, & Bi, 2009), although there is some evidence that variability does not influence, or even hampers, learning (Bergmann & Cristia, 2018; Jusczyk, Pisoni, & Mullennix, 1992). Most such studies have introduced variability through additional stimuli, for example by presenting stimuli that exhibit learnable regularities along with stimuli that are unpredictable. One example showing clear benefits of such variability is a language learning study (Gómez, 2002) in which adults and infants heard three-word phrases in an artificial language. In the artificial language, there was a deterministic relationship between the first and third words in the phrase, but the second word was drawn randomly from a set of possible second words. Adults and infants who heard a language with a larger set of second words (i.e. the variability within phrases was larger) learnt the relationship much better than those who had heard less variable phrases because their language had a smaller set. The authors conclude that this variability helps learners because it allows them to zoom out from specific exemplars and find the abstract regularity in the input.

This type of variability has also been introduced as changes in the visual environment. In a statistical learning task, 8-month-old infants directed their attention more to learnable statistical regularities during the intermittent presence of a distractor in the environment (Tummeltshammer & Kirkham, 2013). The infants looked longer to events with a deterministic relationship when an unrelated part of the stimulus space changed colour at unpredictable moments than when events with the same deterministic relationship happened without other changes in the display. In another study, Kidd and colleagues (2012) showed that when infants view sequences of events that could be more or less unexpected, they are more likely to look away from the display when an event is either very expected or very unexpected. The authors suggest that this preference for moderately surprising events is driven by infants making a trade-off between the value of learning new information and the cost of allocating too many cognitive resources (Kidd et al., 2012). Moderate amounts of noise may be another way to indicate to infants that there is learnable structure in the environment, while not demanding excessive cognitive resources to identify that structure.

While these previous results are informative, most designs measure the impact of distractor signals on the learner. These distractors are not along the dimension of interest, and as such are examples of environmental variability, as opposed to noise. We therefore do not yet know how participants learn a relationship when there is a noisy stimulus from which they must use perceptual inference to extract the meaningful signal. The distinction between learning in a variable context and learning from a noisy signal is an important one for understanding the mechanisms involved. Learning in a variable context requires inhibition of irrelevant information, for example if participants never listened to the second word in the phrases used by Gómez (2002), they could learn the relationship perfectly well. Learning from a noisy signal, however, requires the incoming information to be segmented into meaningful signal and superfluous noise which first requires learning how to dissociate the two. This requires tuning of perceptual inference.

During perceptual inference, we use our expectations based on previous experience to interpret what we are sensing. We form a percept by combining those expectations with the incoming information, and the weight given to each is determined by its certainty. When an observer is first learning about possible events, they are relatively unbiased because they have very few expectations or their expectations are very uncertain. This leads to very little weight being given to the expectation, and perception is mostly bottom-up and stimulus driven. As the observer gains more experience, however, their expectations become stronger and begin to bias perception more, which is generally beneficial to learning since it incorporates previous knowledge gains. Conversely, when the signal is noisier, we are less certain of our sensory input so it weighs less against our expectations. When a signal is cleaner, we are more certain of the sensory input so it weighs more against our expectations. This means that the process of perceptual inference theoretically inherently incorporates an estimate of what is signal and what is noise.

As well as certainty, the influence of expectations on perception can also be influenced by the level of detail of the expectation (Kwisthout, Bekkering & van Rooij, 2017). For example, in the case of language acquisition, the phenomenon of perceptual narrowing shows that as infants learn more about their native language, they begin to form stronger expectations that guide their perception and as a consequence they no longer distinguish between similar phonemes that are not used in their native language. These expectations are tuned to a level of detail that is useful for learning their own language, and they no

longer make fine-grain distinctions between sounds that are functionally equivalent in their native language. In this instance, making a more precise expectation would actually hamper learning; tracking every articulation difference between sounds that are functionally equivalent is resource-intensive and would give a worse estimate of the structure of the language's sound system. To learn from new contexts, it is therefore necessary to start to abstract from the small fluctuations in the environment, the noise, and try to identify the statistical regularities, the signal.

This weighing of expectations and sensory input, and its tuning using certainty and level of detail, is happening constantly and underlies all further information processing. As such, it is fundamental to cognition, but it has largely been overlooked in developmental science. Current theories do not account for the development of this process and instead posit a static adult model (Kayhan & Kwisthout, 2017). We therefore ask how infants respond to a noisy signal, specifically how fast they learn a pattern with different levels of noise added. We believe that using online learning measures to examine the time-course of the effects of this type of noise will reveal crucial insight into how infants make sense of their sensory input.

In the current study, we manipulate information along the dimension of interest, that is, the location of a stimulus, such that there are also small fluctuations in its location at each appearance. Crucially, infants must learn the correct level of detail for their expectation to identify the statistical regularity. To learn the pattern, the expectation "the stimulus will appear somewhere in region A" is good enough, and the specific pixels the stimulus will appear at will be uninformative and hinder pattern recognition.

Not only will this paradigm show how infants separate noise from signal, it will also show how infants learn from events that are more or less unexpected. Moderately unexpected events hold infants' attention longer than stimuli that are too expected or too unexpected, suggesting that infants find moderately surprising stimuli the most informative (Kidd, Piantadosi & Aslin, 2012; Kidd, Piantadosi, & Aslin, 2014). Following from this study of preference for intermediately unexpected events (Kidd et al., 2012), we ask whether this preference serves a function. That is, can infants learn faster from a sequence of events that are intermediately unexpected? Due to the design of the task of Kidd and colleagues (2012), there was no opportunity for the infants to demonstrate learning, as the sequences did not repeat. While the infants did show evidence of accumulating knowledge about the

experimental environment by looking away when an event was too expected or too unexpected, there was no way to learn which stimulus was upcoming, since the events were independent from one another. In the current paradigm, it is possible to manipulate how close the stimulus appears to the expected location, if expectations are built on past experience as we would expect from perceptual inference accounts. Infants can then demonstrate what they have learnt with their eye movements, giving a measure of how much they are learning and how this is affected by the expectedness of the stimulus location.

The Current Study

In the current study we expand upon previous studies of variability by examining the time-course of infants' learning of a sequence of deterministic events with and without noise in the realisation of these events. Understanding how infants learn from noisy signals will ultimately allow us to build ecologically valid models of how babies learn in the real world. Since all sensory input comes as a package of information consisting of signal and noise, and the amount of noise is much larger in real-world learning than in the lab, it is crucial that we understand how infants identify and distinguish between signal and noise if we want to know how learning really happens.

Study Design

To investigate this, we will test 7- and 8-month-old infants, manipulating variability within-participants using an implicit learning paradigm, in which the infants are expected to make saccades more quickly as they learn a pattern of events (see statistical learning literature, e.g. Haith, 1993; Kirkham, Slemmer, Richardson, & Johnson, 2007). We will track the infants' eye movements in order to measure their learning rates and visual attention. In our novel task, a frog appears for 1500ms on one of four leaves in a deterministic pattern which is repeated 10 times without breaks. Each infant sees three blocks, each with a different sequence, one of which is presented with no noise, one with moderate noise and one with high noise (see Figure 1, and for gifs see https://osf.io/f4mhc/?view_only=b718a876674e419b81b6df5c304f4ad5). The patterns each consist of five frog appearances, and are created pseudorandomly such that there are the same number of diagonal movements in each sequence and the frog cannot appear twice consecutively on the same leaf. The final position of the pattern also cannot be the same as the starting position, since the sequence repeats without breaks and this would be in fact

constitute appearing twice consecutively on the same leaf. One sequence per block is selected randomly at the time of stimulus presentation such that sequence is entirely independent of noise level, and no infant will see the same sequence in more than one condition.

Figure 1: Trial schematic showing example sequences in the (A) no-variability condition, (B) moderate-variability condition and (C) high-variability condition. The frog appears in only one position at a time, and moves from one leaf to the next in the pattern shown by the arrows. Each sequence is five items long, and is repeated in a continuous stream.

In the no-noise condition, the frog always appears in the centre of the leaf, and the sequence is visually identical on every iteration. In the moderate-noise and high-noise conditions, the frog still appears on the leaf in a deterministic sequence, but the exact position on the leaf is randomly jittered. This leads to events with a learnable structure (if the frog is on leaf 1, it will always appear next on leaf 2), but which are distinct from each other in each instantiation due to the jitter. The frog may appear on the top left of leaf 1 in the first repetition of the pattern and the bottom centre of leaf 1 in the next repetition, but the consequences will be the same: the frog will next move to somewhere on leaf 2.

Hypotheses

The goal of the current experiment is to examine how infants tune the level of detail of their expectations during perceptual inference. In order to address this question we ask whether noise affects learning rate in infants in an implicit sequence-learning task and to measure the effect of levels of noise on learning rates. Since variable contexts seem to highlight relevant information (e.g., Gómez, 2002; Wulf & Schmidt, 1997), and noise is a specific type of variability along the dimension of interest, we propose that increased noise may also direct attention to relevant information and lead to faster learning. We

hypothesise that the infants will learn the sequence at all levels of noise, but that they will learn faster when there is some noise than when there is none. We also hypothesise that infants will learn fastest in the moderate-noise condition, and that this condition will hold their visual attention the longest.

Measures

In order to quantify the effect of noise on learning, we will analyse the infant's learning rates, calculated from their saccadic reaction times (Desrochers, Kurdziel, & Spencer, 2016; Thomas & Nelson, 2001; Tummeltshammer & Kirkham, 2013). We expect that infants will learn the sequence in all 3 conditions, but that they will learn fastest in the moderate-noise condition, i.e., there will be an interaction between number of repetitions and condition, with the largest decrease in reaction times over repetitions in the moderate-noise condition. We also expect that if the infants learn the sequence fully, their reaction times will plateau, and the time to plateau will be shorter in the moderate-noise condition than in the no-noise and high-noise conditions.

To test our secondary hypothesis, we will measure the visual attention of the infants to the display, and compare total looking time and time until first lookaway between the different levels of noise (cf. Kidd et al., 2012). We expect the infants to look away from the screen earlier in the no-noise and high-noise conditions, and to attend to the screen less overall during those conditions, than in the moderate-noise condition.

Method

Participants

Participants will be infants between the ages of 7 months 0 days and 8 months 30 days. They will be recruited through a database of volunteer families. Informed consent will be received from one parent for each infant, and the infants will receive a small thank-you present of 10 euros or a book for participation. Infants will be excluded from participating if they were born before 36 weeks of gestation, or if they have known vision impairments or neurological problems at the time of testing.

We will test enough infants to acquire 36 datasets that meet our inclusion criteria detailed in the Data Analysis section below. This sample size is based on a power calculation using a bespoke web application (Westfall, Kenny & Judd, 2014). We computed the sample size needed to detect an effect size of Cohen's $d=0.8$, approximately equal to a partial η^2 of

0.12, based on typical condition differences in similar designs (Tummeltshammer & Kirkham, 2013; Tummeltshammer, Mareschal & Kirkham, 2014; Tummeltshammer & Amso, 2016). To achieve 90% power, our planned model structure requires 32.4 participants, assuming that parameters explain the same amount of variance as they did in Tummeltshammer and Kirkham's (2013) model. We have decided to acquire 36 analysable datasets to be conservative, and leave room for potential differences between our design and the previous studies.

Apparatus

Infants' eye movements will be recorded using a Tobii T120 eye-tracker with an integrated monitor. Stimuli will be presented using Psychtoolbox (version 3.0.11) and Talk2Tobii toolbox within Matlab (version 2013b), and sounds will be played through external stereo speakers. The infant will be monitored through an external video camera. Audiovisual attention-getters will be played between blocks to keep the infants engaged and to minimise experimenter influence on look-aways.

Design and Procedure

All infants will be tested with a parent in a quiet room, seated in a highchair approximately 60 cm away from the screen. A 9-point calibration procedure will be used to obtain a reliable signal. Infants will be required to fixate each point before the experimenter manually advances the calibration sequence and the sequence will be repeated if fewer than 5 points (centre and 4 corners) were accurately calibrated.

After calibration, the infants will be presented with the implicit learning task, consisting of three blocks. Level of variability (none, moderate, high) will be counterbalanced between participants, and each block will have a pseudorandom sequence. Stimulus presentation will last approximately 5 minutes and the session will be terminated at parent request; if the infant stops attending to the screen despite the use of attention getters and experimenter encouragement; or if the infant cries or otherwise shows distress.

Stimuli

Stimuli will consist of a fixed array of circular lilypads and an image of a frog, which appears on each lilypad according to a sequence which will be repeated within each block. In the no-noise block, the frog will appear in the centre of the leaf. In blocks with added noise, the frog's position on each leaf is jittered from the leaf centre by a number of pixels randomly selected from the uniform distribution over the circle. This jitter is generated

separately for every trial, which makes it neither predictive nor predictable. In the moderate-noise condition, the maximum jitter will be 50% of the leaf radius, and in the high-noise condition, the maximum jitter will be 100% of the leaf radius.

Data Analysis

Datasets will only be included for analysis if the participant sees a sufficient number of events. Initially, our criterion will be to include blocks from a particular infant in the final dataset only if the infant saw more than half of the events in that block, and to include infants in the final sample only if they meet this criterion for at least two of the three blocks. If this leads to a dropout rate of more than 25% of the infants, we will reduce the criterion to 40% of events. We will segment all remaining datasets into fixations using the Identification by 2-Means Clustering method, with the recommended values of 40ms minimum fixation duration and merging fixations within 0.7° of each other (Hessels, Niehorster, Kemner & Hooze, 2017). We will define square areas of interest around each leaf in order to group fixations to each leaf. After processing, datasets will only be included in the final analysis if the learning curve calculated from fixation-segmented data in at least one included condition is different from zero, unless this also leads to a dropout rate of more than 25% of the infants. If a large proportion of tested infants show no learning curve, this is informative for our hypotheses, and will be taken into account in the discussion.

Events to which the infant's response was slower than 2 standard deviations below their individual median response time will also be excluded. To ensure that we include only looks that indicate sequence learning, we will also exclude incorrect looks. Incorrect looks will be defined as looks not within the region of interest around a leaf where the frog was currently present and not within the region of interest around a leaf where the frog was about to appear in the next 500ms. We include this 500ms period before the frog's arrival to capture anticipatory looks to the upcoming position.

We will measure the latency of infants' eye movements to the frog's new position as well as looking time to the display comparable to the measures used by Tummeltshammer and Kirkham (2013). Using the latencies we will compute a learning curve, which is defined as median saccadic latency over positions for each repetition of the sequence (Desrochers et al., 2016; Thomas & Nelson, 2001), for descriptive purposes and data visualisation. We will then analyse the raw data using multi-level modelling with random intercept for participants, random intercept for stimuli, and random slope of repetitions and condition by

participant. We expect that participants will have random intercepts due to individual differences in response times, and random slope of repetitions and condition by participant due to individual differences in learning rates. We model fixed effects of repetition and condition and their interaction, in order to capture the progression of learning throughout each block. We expect that the slope of the learning curve will be steepest in the moderate noise condition, since we expect that moderate noise should facilitate learning.

Using the looking times to the display, we will calculate how long infants attend to the screen in total, as measured by total fixation duration, and how long until infants' first look-away of more than one second, comparable to the procedure of Kidd and colleagues (2012). This will give one value per participant per condition for each measure and we expect that both measures will form an inverted u-shape, with longest total looking time and longest time to first look-away in the moderate noise condition. We will then run a regression model for each measure, to fit a quadratic curve to the data.

Model Specifications

Before running the models, we will test the raw data for normality, skew, kurtosis and multicollinearity, and perform any necessary transformations based on the results (Field, Miles & Field, 2012). We will then build the models incrementally, starting from the most parsimonious model and including the *a priori* terms listed above one at a time. Fixed effects will be retained in the final model if they have an associated *t* value of larger than 1.96, and random effects will be retained in the final model if this decreases the deviance significantly based on the Restricted Maximum Likelihood. The final models will be compared using the Bayesian Information Criterion and the model with the lowest score will be selected.

The models that we *a priori* built according to our hypotheses are as follows:

```
model_learning_rate <- lmer (RTlog ~ Condition * Repetition + Item + (1 + Condition | Participant))
```

```
model_total_fix_duration <- lmer (total_fix_duration ~ Condition2 (1 + Condition | Participant))
```

```
model_first_lookaway <- lmer (time_to_first_lookaway ~ Condition2 (1 + Condition | Participant))
```

DOES VARIABILITY HELP INFANTS LEARN?

and will be run in R version 3.5.1 (R Core Team, 2018) using R Studio version 1.1.456 (RStudio Team, 2016), with the following packages: arm (Gelman & Su, 2018), lattice (Sarkar, 2008), lme4 (Bates, Maechler, Bolker & Walker, 2015), moments (Komsta & Novomestky, 2015), mvoutlier (Filzmoser & Gschwandtner, 2018), and psych (Revelle, 2018).

References

- Aggelopoulos, N. C. (2015). Perceptual Inference. *Neuroscience and Biobehavioral Reviews*, 55, 375—392. <http://dx.doi.org/10.1016/j.neubiorev.2015.05.001>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. <https://doi.org/10.18637/jss.v067.i01>
- Bergmann, C., Cristia, A., & Dupoux, E. (2016). Discriminability of sound contrasts in the face of speaker variation quantified. 38th Annual Meeting of the Cognitive Science Society (CogSci), on August 11-13.
- Brainard, D. H. (1997) The Psychophysics Toolbox, *Spatial Vision* 10:433-436.
- Clark, A. (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204.
- Desrochers, P. C., Kurdziel, L. B. F., & Spencer, R. M. C. (2016). Delayed benefit of naps on motor learning in preschool children. *Experimental Brain Research*, 234(3), 763–772. <https://doi.org/10.1007/s00221-015-4506-3>
- Douvis, S. J. (2005). Variable practice in learning the forehand drive in tennis. *Perceptual and motor skills*, 101(2), 531—545.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.
- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London: Sage publications.

Filzmoser, P. and Gschwandtner, M. (2018). mvoutlier: Multivariate Outlier Detection Based on Robust Methods. R package version 2.0.9. <https://CRAN.R-project.org/package=mvoutlier>

Gelman, A. and Su, Y.-S. (2018). arm: Data Analysis Using Regression and Multilevel/Hierarchical Models. R package version 1.10-1. <https://CRAN.R-project.org/package=arm>

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13(5), 431—436. <https://doi.org/10.7897/2277-4343.06112>

Gonzalez, C., & Madhavan, P. (2011). Diversity during training enhances detection of novel stimuli. *Journal of Cognitive Psychology*, 23(3), 342–350. <https://doi.org/10.1080/20445911.2011.507187>

Haith, M. M. (1993). Future-oriented processes in infancy: The case of visual expectations. In Granrud, C. E. (Ed.) *Visual perception and cognition in infancy*, 235—264. New Jersey, USA: Lawrence Erlbaum Associates.

Hessels, R. S., Niehorster, D. C., Kemner, C., & Hooge, I. T. (2017). Noise-robust fixation detection in eye movement data: Identification by two-means clustering (I2MC). *Behavior research methods*, 49(5), 1802-1823. <https://doi.org/10.3758/s13428-016-0822-1>

Jusczyk, P. W., Pisoni, D. B., & Mullennix, J. (1992). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition*, 43(3), 253-291.

Kayhan, E. & Kwisthout, J. (2017). Predictive Processing in Development. *The Newsletter of the Technical Committee on Cognitive and Developmental Systems*, 14(1). Retrieved from <http://www.cse.msu.edu/amdtc/amdnl/CDSNL-V14-N1.pdf>

DOES VARIABILITY HELP INFANTS LEARN?

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE*, 7(5), 1–8. <https://doi.org/10.1371/journal.pone.0036399>

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The goldilocks effect in infant auditory attention. *Child Development*, 85(5), 1795–1804. <https://doi.org/10.1111/cdev.12263>

Kirkham, N. Z., Slemmer, J. A., Richardson, D. C., & Johnson, S. P. (2007). Location, location, location: Development of spatiotemporal sequence learning in infancy. *Child Development*, 78(5), 1559–1571. <https://doi.org/10.1111/j.1467-8624.2007.01083.x>

Komsta, L. & Novomestky, F. (2015). moments: Moments, cumulants, skewness, kurtosis and related tests. R package version 0.14. <https://CRAN.R-project.org/package=moments>

Kwisthout, J., Bekkering, H., & Van Rooij, I. (2017). To be precise, the details don't matter: on predictive processing, precision, and level of detail of predictions. *Brain and Cognition*, 112, 84-91.

Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition*, 131(2), 284-299.

RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. <http://www.rstudio.com/>.

Reddy, V., Markova, G., & Wallot, S. (2013). Anticipatory adjustments to being picked up in infancy. *PloS one*, 8(6), e65289.

Revelle, W. (2018) psych: Procedures for Personality and Psychological Research. Version=1.8.12. Northwestern University, Evanston, Illinois, USA. <https://CRAN.R-project.org/package=psych>

- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, 8(2), 101–105.
- Sarkar, D. (2008) Lattice: Multivariate Data Visualization with R. Springer, New York. ISBN 978-0-387-75968-5. <http://lmdvr.r-forge.r-project.org>
- Thomas, K. M., & Nelson, C. A. (2001). Serial reaction time learning in preschool-and school-age children. *Journal of experimental child psychology*, 79(4), 364-387.
- Tummeltshammer, K., & Amso, D. (2017). Top-down contextual knowledge guides visual attention in infancy. *Developmental Science*, (November 2016), 1–9. <https://doi.org/10.1111/desc.12599>
- Tummeltshammer, K. S., & Kirkham, N. Z. (2013). Learning to look: Probabilistic variation and noise guide infants' eye movements. *Developmental Science*, 16(5), 760–771. <https://doi.org/10.1111/desc.12064>
- Tummeltshammer, K. S., Mareschal, D., & Kirkham, N. Z. (2014). Infants' selective attention to reliable visual cues in the presence of salient distractors. *Child development*, 85(5), 1981-1994. <https://doi.org/10.1111/cdev.12239>
- Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, 143(5), 2020-2045.
- Wulf, G., & Schmidt, R. A. (1997). Variability of practice and implicit motor learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23(4), 987–1006.

DOES VARIABILITY HELP INFANTS LEARN?

Yao, W. X., DeSola, W., & Bi, Z. C. (2009). Variable practice versus constant practice in the acquisition of wheelchair propulsive speeds. *Perceptual and motor skills*, 109(1), 133-139.

DOES VARIABILITY HELP INFANTS LEARN?

Figure 1: Trial schematic showing example sequences in the (A) no-variability condition, (B) moderate-variability condition and (C) high-variability condition. The frog appears in only one position at a time, and moves from one leaf to the next in the pattern shown by the arrows. Each sequence is five items long, and is repeated in a continuous stream.