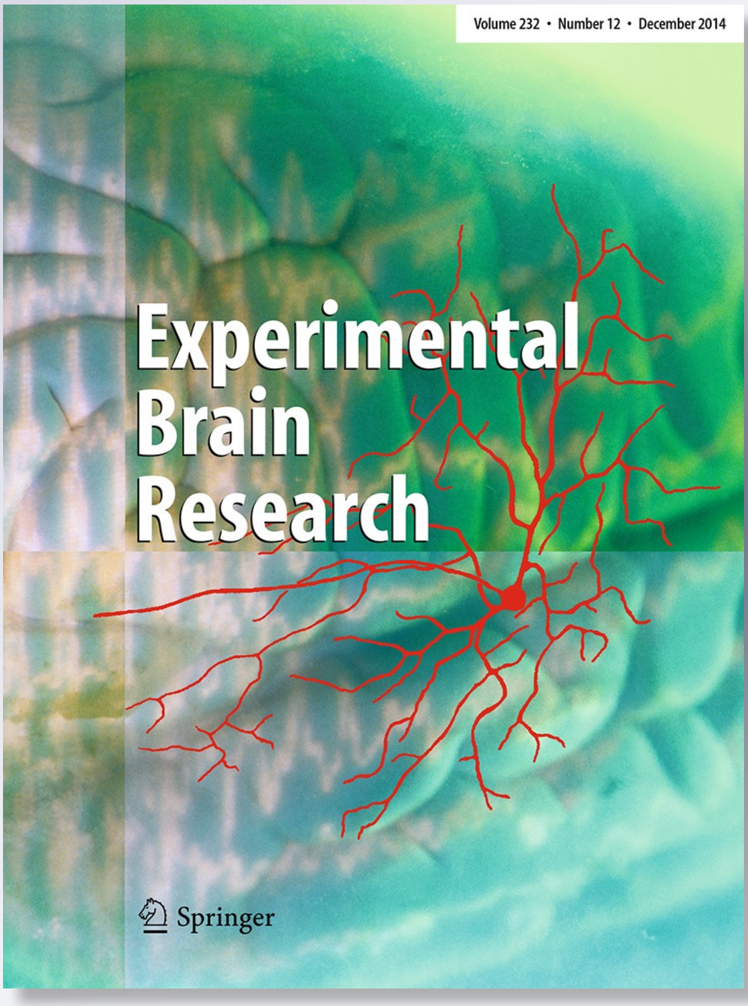
Development of context dependency in human space perception

Alessandra Sciutti, David Burr, Alice Saracco, Giulio Sandini & Monica Gori

Experimental Brain Research

ISSN 0014-4819
Volume 232
Number 12

Exp Brain Res (2014) 232:3965-3976
DOI 10.1007/s00221-014-4021-y



Springer

Your article is protected by copyright and all rights are held exclusively by Springer-Verlag Berlin Heidelberg. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com”.

Author's personal copy

Exp Brain Res (2014) 232:3965–3976
DOI 10.1007/s00221-014-4021-y

RESEARCH ARTICLE

Development of context dependency in human space perception

Alessandra Sciutti · David Burr · Alice Saracco ·
Giulio Sandini · Monica Gori

Received: 31 December 2013 / Accepted: 13 June 2014 / Published online: 3 September 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract Perception is a complex process, where prior knowledge exerts a fundamental influence over what we see. The use of priors is at the basis of the well-known phenomenon of central tendency: Judgments of almost all quantities (such as length, duration, and number) tend to gravitate toward their mean magnitude. Although such context dependency is universal in adult perceptual judgments, how it develops with age remains unknown. We asked children from 7 to 14 years of age and adults to reproduce lengths of stimuli drawn from different distributions and evaluated whether judgments were influenced by stimulus context. All participants reproduced the presented length differently depending on the context: The same stimulus was reproduced as shorter, when on average stimuli were short, and as longer, when on average stimuli were long. Interestingly, the relative importance given to the current sensory signal and to priors was almost constant during childhood. This strategy, which in adults is optimal in Bayesian terms, is apparently successful in holding the sensory noise at bay even during development. Hence, the influence of previous knowledge on perception is present already in young children, suggesting that context dependency is established early in the developing brain.

Keywords Length perception · Development · Vision · Perceptual priors · Bayesian modeling

A. Sciutti (✉) · A. Saracco · G. Sandini · M. Gori
Robotics, Brain and Cognitive Sciences Department, Istituto Italiano di Tecnologia, Via Morego 30, 16163 Genoa, Italy
e-mail: alessandra.sciutti@iit.it

D. Burr
Department of Psychology, University of Florence, Florence, Italy

Introduction

Our perception of the world depends not only on the incoming sensory information, but it is also influenced by our prior knowledge. This phenomenon traces back to Helmholtz, who introduced the concept of unconscious inference, which helps to shape vision (Helmholtz 1866). Incorporating previous experience, or *priors*, into the current percept helps the brain cope with the uncertainty resulting from sensory and neural noise, and ambiguity (e.g., the mapping of 3D objects into two-dimensional images, Knill and Pouget 2004). The use of priors improves the reliability of perception, reducing the overall noise. It is often considered to reflect a statistically optimal computation, which can be modeled by Bayesian techniques (Weiss et al. 2002; Kersten and Yuille 2003).

Use of priors has been recently suggested to be at the basis of the well-known phenomenon of central tendency (Jazayeri and Shadlen 2010). Judgments of almost all quantities (such as length, duration, number, and color) tend to gravitate toward their mean magnitude, thus implying that on any range of values the lower end of the range will be overestimated and the higher end will be underestimated (Hollingworth 1910). These biases in magnitude estimation occur with everybody and become more exaggerated in neglect patients in a wide array of perceptual continua and across multiple modalities (visual, auditory, proprioceptive, tactile and even gustatory, Mennemeier et al. 2005). Recent works indicate that this phenomenon can be modeled within a Bayesian framework, where the mean magnitude of the stimulus history can be described as a prior, biasing the judgments toward the average (Jazayeri and Shadlen 2010; Cicchini et al. 2012). These authors, who studied the reproduction of

Author's personal copy

time intervals, showed that the estimate of a sample duration differed depending on the distribution from which it was drawn; so the same time interval was reproduced as longer if on average stimuli lasted longer, and as shorter if on average previous stimuli were shorter (Jazayeri and Shadlen 2010). Single judgments were therefore inaccurate, as they did not reproduce exactly the current stimulus duration. Nonetheless, the increased reliability associated with this strategy led to an overall reduction in the error magnitude, showing that the use of priors can serve as a filter, smoothing variations in the sensory input, which are often due to noise and hinder stable perceptions.

Although central tendency is universal in adult perceptual judgments, how this strategy develops with age remains unknown. While many sensory aspects develop early, within the first years of age (e.g., Atkinson 2000; Streri 2003), many others develop later in life, during school years (e.g., Morrongiello et al. 1994; Rentschler et al. 2004; Gori et al. 2008, 2012a). Therefore, a strategy aimed at minimizing total error in sensory judgments could be particularly beneficial for children. On the other hand, for the developing brain, it could be fundamental to formulate estimations as veridical as possible, so that—by trial and the feedback of the error—children could develop the ability to produce accurate judgments. Thus, a Bayesian model aimed at minimizing the total error at the expense of accuracy, although theoretically optimal in reducing error, may be detrimental in childhood. For example, cross-sensory fusion based on statistically optimal integration strategies develops late (at 8–10 years of age). Before then, children base their estimates on one sense rather than on the fused output. The authors speculate that this may not be the more precise, but the more robust (Gori et al. 2008). This strategy—which is sub-optimal in Bayesian terms—has been suggested to be useful for cross-sensory calibration (Burr and Gori 2011).

In this study, we measured the development of central tendency in a spatial task by assessing precision and accuracy of the reproduction of a perceived length. We asked children ranging from 7 to 14 years of age and adults to reproduce lengths drawn from different distributions and evaluated whether judgments were influenced by previous stimulus history. We separately estimated the precision of length discrimination, to assess whether the central tendency strategy correlated with sensory precision during development. All the data were modeled within a Bayesian context, using techniques similar to those of Cicchini et al. (2012). The analysis was conducted to ascertain which relative weighting of prior and sensory input was consistent with the data, if observers used an optimal reliability-based weighting strategy.

Materials and Methods

Subjects

Eighty-two children aged 7–14 years from elementary and middle schools in Genoa (Italy) and 6 adults, recruited from the local university took part in the study. Nineteen participants were excluded from analysis (eight 7-year-olds, seven 8-year-olds, three 10-year-olds and one 11-year old) on the basis of two criteria: a spatial discrimination threshold larger than 75 % of maximum variation of stimulus length tested in a single session (i.e., spatial discrimination threshold measured in the *Length discrimination* over 6 cm) or the inability to perform the task. This latter condition was verified when one participant reproduced with <5 cm difference the two extreme amplitudes presented in the experiment—2 and 14 cm—or when the regression of reproduced lengths over stimulus lengths exhibited a negative slope. This high selection criterion was chosen to avoid the inclusion in the sample of subjects who showed difficulty in performing the task, or in maintaining attention for the relatively long periods required.

The remaining sample therefore comprised 65 children: 7 seven-year-olds (6.7 ± 0.5 years, mean ± SD), 9 eight-year-olds (7.9 ± 0.6 years), 15 ten-year-olds (9.9 ± 0.4 years), 18 eleven-year-olds (11.4 ± 0.6 years), 16 thirteen-year-olds (13.3 ± 0.6 years) and 6 adults (29.8 ± 7.1 years). Children were compensated with a €10 bookstore coupon for their participation. All subjects had normal or corrected to normal visual acuity. All participants gave written informed consent prior to testing, and the study was approved by the local ethics committee (Azienda Sanitaria Locale Genovese N.3).

Stimuli

The experiments were performed in a dimly lit room. Visual stimuli were presented on a LG FLATRON L1730 SF Touchscreen (screen dimension: 398 × 402 mm; screen resolution: 1,280 × 1,024 pixels). Subjects were seated in front of the screen, with their head aligned with screen center and with the shoulder at a distance corresponding to 60 % of their extended arm length. This position was chosen to allow each participant to comfortably reach the touch screen. The visual stimuli were created with the MATLAB Psychophysics Toolbox (Brainard 1997; Kleiner et al. 2007). Subject responses were collected through the touch screen connected via USB to the laptop.

Procedure

The experiment comprised two different tasks: *length reproduction*, measuring central tendency in estimates of

Author's personal copy

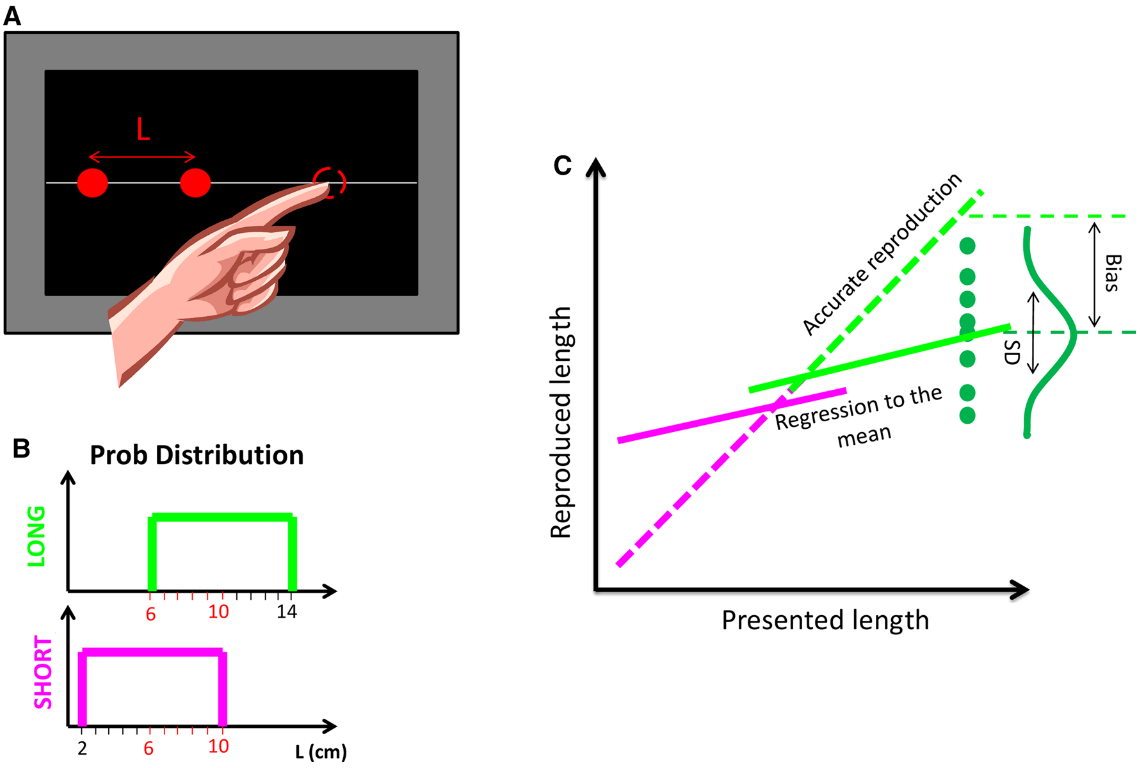


Fig. 1 Experimental methods. **a** Sketch of the *Length reproduction* experimental procedure: On each trial, subjects were presented with two subsequent flashes of light positioned along a visible straight white line crossing the whole screen at its middle height. After the disappearance of the second flash, subjects were requested to touch a point on the straight line, so that the distance between the touched point from the second disk of the stimulation were equal to the distance between the first and the second disk. **b** Probability distribution of stimulus lengths (L) in the *Long* and *Short* experimental sessions. **c** Schema of the possible results of the *Length reproduction* task. For an accurate reproduction, length estimates would lie on the *bisector line* (*identity line*) and reproductions of the same stimulus length belonging to the two different contexts (Long and Short) would be superimposed. A central tendency would imply a flatter line describing the regression of estimates over presented lengths, and a difference in the reproduction of a stimulus as a function of stimulus context. The Gaussian on the right represents a distribution of reproductions for a single stimulus length. The average error in the reproduction is partitioned into Bias (difference between the average reproduction and real stimulus length) and variability (SD, standard deviation of all the reproductions for that stimulus). Both indicators will be divided by average stimulus length (see "Methods" section for more details)

lengths at different ages; *length discrimination*, designed to measure the perceptual precision of length judgments.

Length reproduction

On each trial, subjects were presented with two consecutive light flashes (red disks of 1 cm diameter, duration 400 ms) positioned along a visible straight white line crossing the whole screen at its central height. The first flash was positioned at a variable distance from the left border of the screen (0.5–3.5 cm, randomly selected). On its disappearance, a second disk appeared at a variable distance to the right of it, disappearing after 400 ms. Subjects were requested to touch a point on the straight line in order to reproduce the distance between the first and the second disk of the stimulation (Fig. 1a). After the pressure, a red disk came on to indicate where the subject had touched, but no feedback about the correctness of the response was provided. Each new trial started on the experimenter's button press, with the first light appearing after 500 ms. Each subject participated in two sessions: a Short condition, in which the spatial distance between the two first disks ranged from 0.2 to 10 cm, and a Long condition, in which the distances presented ranged from 6 to 14 cm (Fig. 1b). The order of the sessions was randomized between subjects. Each session was characterized by 11 different sample intervals (separated by 0.8 cm each), each of which was presented 7 times, yielding to a total of 77 trials per subject per condition. The duration of each session was about 30 min (1 h for both conditions). We chose to keep the order of disk appearance constant to keep the task as simple as possible for the youngest age group. The Short and Long sessions were measured for each subject in two separate days, to avoid any influence of the exposure to different stimulus distributions in the same day.

Author's personal copy

Length discrimination

A length discrimination task was designed to evaluate the precision of visual space perception. The visual stimuli were the same as those used in the length reproduction task: Red disks of 1 cm diameter positioned along a white straight line crossing the screen at its central height. The experimenter initiated each trial by pressing a button. After 500 ms, three red disks were presented simultaneously for 400 ms in different positions along the line. After stimulus disappearance, subjects had to judge whether the second disk was closer to the first or the third disk. One of the two distances—chosen randomly on each trial—maintained always the same length (standard), while the other—the comparison—changed from trial to trial according to a QUEST adaptive procedure [starting value: 9 ± 1.8 cm (SD), step size: 2 mm; range 16.2 cm; β: 3.5; δ: 0.01; γ: 0; threshold criterion: 0.5; additional noise: 20 %: Watson and Pelli (1983)]. This design was selected after some piloting because it represented a very simple measure of length discrimination, where priors could not influence performance. The proportion of times in which the comparison interval was judged longer than the standard was plotted as a function of comparison amplitude and fit by a cumulative Gaussian distribution. Mean and standard deviation of the fit provided the point of subjective equality (PSE) and a bias-free estimate of precision, respectively. Standard errors for the PSE and the precisions were obtained by bootstrapping (Efron and Tibshirani 1993). In this task, we also ran two separate conditions in separate sessions: the Short condition, with the standard distance fixed at 6 cm (the average of the stimulus length presented in the Short reproduction session), and a Long condition, with the standard interval fixed at 10 cm (corresponding to the average stimulus interval of the LONG reproduction task). Forty trials per condition were collected for each subject. The Weber fraction was computed for each session as the ratio between the measured precision and the average length of the stimuli in that session.

Data analysis

As in Cicchini et al. (2012) and Jazayeri and Shadlen (2010), we partitioned the total error of the reproduction tasks into two parts: the bias, corresponding to the accuracy of the reproduction, and the coefficient of variation (CV: standard deviation divided by the average stimulus length), indicating the precision of the reproduction (see SD in Fig. 1c). To compute these two indicators, we first subtracted from each reproduced distance the average reproduced distance for all the trials ($\bar{R}_c$) and summed the central stimulus distance $\bar{S}$, to remove any constant biases (rather than regression toward the mean):

$$R'_{i,n} = R_{i,n} - \bar{R}_c + \bar{S} \tag{1}$$

where $R_{i,n}$ represents the reproduced distance for the n-th repetition of the interval stimulus i.

The bias, for each i-th stimulus value, corresponds to the difference between the average produced distance and the sample distance, normalized by the average length of the stimuli in a session:

$$BIAS_i = \frac{\left|\bar{R}_i - S_i\right|}{\bar{S}}. \tag{2}$$

CV is given by the standard deviation of the N reproduced distances for each i-th stimulus distance, again normalized by average stimulus distance in that session:

$$CV_i = \frac{\sqrt{\sum \left(R'_i - \bar{R}'_i\right)^2 / N}}{\bar{S}}. \tag{3}$$

The total error for each i-th sample stimulus is given by the Pythagorean sum of bias and CV for that stimulus:

$$RMSE_i = \sqrt{BIAS_i^2 + CV_i^2}. \tag{4}$$

As a direct measure of central tendency, the regression index was computed as the difference in slope between the best linear fit of the reproduced lengths and the identity line. This index varies from 0 (veridical performance) to 1 (complete regression to the mean). Regression index values were tested for significance with respect to 0 with one-sample t tests. To assess potential developmental changes in perceptual strategies, all performance parameters have been submitted to one-way ANOVAs, followed by Bonferroni post hoc tests. The analyses were computed after averaging, for each subject, the results from the Short and the Long sessions.

Bayesian modeling

The central tendency has been modeled as a way to increase the consistency of the perceptual estimations by biasing the response toward the mean of the previously perceived stimuli. This approach reduces the accuracy of the reproduction, but also diminishes the variability associated with the response, a strategy that for noisy sensory estimations (low sensory precision) optimizes reproduction performances, reducing the total error (Jazayeri and Shadlen 2010). This fundamental principle may be modeled in Bayesian terms, where every stimulus reproduction includes an estimate of the statistics of the previously presented stimuli (Jazayeri and Shadlen 2010; Cicchini et al. 2012). The Bayesian model is illustrated in Fig. 2a. Within this framework, to estimate stimulus amplitude, the observer combines two sources of information: the current noisy estimate of stimulus length

Author's personal copy

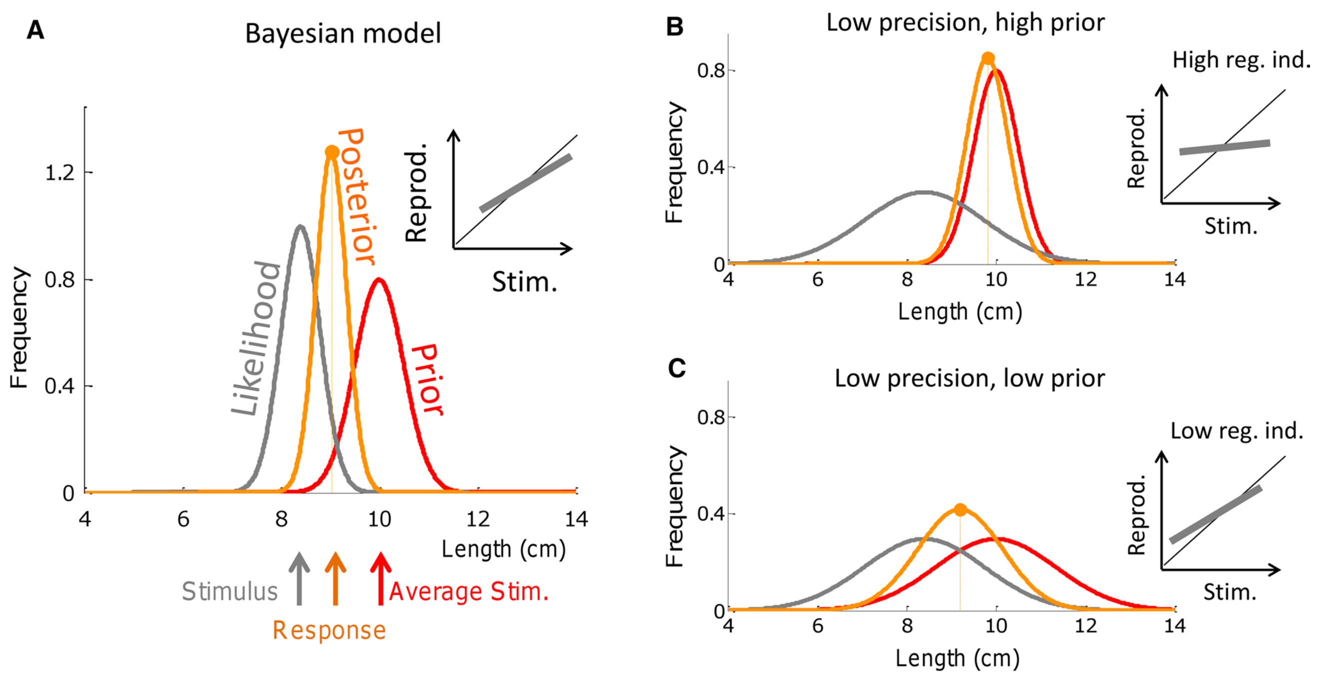


Fig. 2 Illustration of the Bayesian Gaussian prior model (modified from Cicchini et al. (2012)) and its implication of the regression of estimates over real stimulus length (*insets*). The likelihood function for the current stimulus is modeled by a Gaussian centered on the current stimulus and with a width corresponding to subjects' sensory precision, while the prior is represented by a Gaussian probability density function derived from past trials, and hence centered on the average of the stimuli in the session. *Panels* **b** and **c** show two alternative effects of a reduced sensory precision as a function of prior width

and prior knowledge about the stimuli distribution (an approximation of its statistics). In Bayesian terms, these two elements correspond to the *likelihood* function (represented by a Gaussian distribution centered at the current stimulus length, with standard deviation corresponding to the subjects' sensory precision) and the *prior* probability distribution function (also approximated by a Gaussian distribution centered on the average stimulus of the session). According to this model, given a certain prior amplitude, the more participants are precise in their sensory estimate (lower Weber and hence narrower likelihood) the more their reproduction will mimic sensory information. If instead, sensory precision is low (likelihood larger than prior), then reproduction will be more similar to the prior (see Fig. 2b), with a larger regression toward the mean appearing.

In mathematical terms, assuming that both the prior and likelihood function are Gaussians with mean and standard deviations (μ_P, σ_P) and (μ_L, σ_L) and given a stimulus length S_i, the prior will be centered on the average stimulus of that condition $(\mu_P = \bar{S})$ and the likelihood function will be centered on the measurement of the stimulus length $(\mu_L = S_i)$. According to Bayes' rule, the posterior distribution is a Gaussian centered at:

$$\mu_R = \mu_L - \frac{\sigma_L^2(\mu_L - \mu_P)}{\sigma_L^2 + \sigma_P^2} \tag{5}$$

with variance:

$$\sigma_R^2 = \frac{\sigma_L^2 \sigma_P^2}{\sigma_L^2 + \sigma_P^2} = VAR\_OBS \tag{6}$$

which corresponds to the variance of an observer who estimates stimulus length as the maximum of the posterior. Note that σ_R is by construction always smaller than σ_L and σ_P.

Pooling across trials with stimuli of length S_i, the observer bias can be computed as:

$$BIAS\_OBS\_S_i = \frac{\sigma_L^2(S_i - \bar{S})}{\sigma_L^2 + \sigma_P^2}.$$

Which can be extended to a range of spatial stimuli (rather than the single stimulus length) as follows:

$$BIAS\_OBS = \frac{\sigma_L^2 \sqrt{\frac{\sum_i (S_i - \bar{S})^2}{N}}}{\sigma_L^2 + \sigma_P^2}. \tag{7}$$

It is worth noting that this "ideal observer" analysis relies on observers using an optimal (reliability weighted) averaging strategy for sensory and prior information. We cannot tell whether such optimal integration is actually adopted by children, but, assuming this is the case, our

Author's personal copy

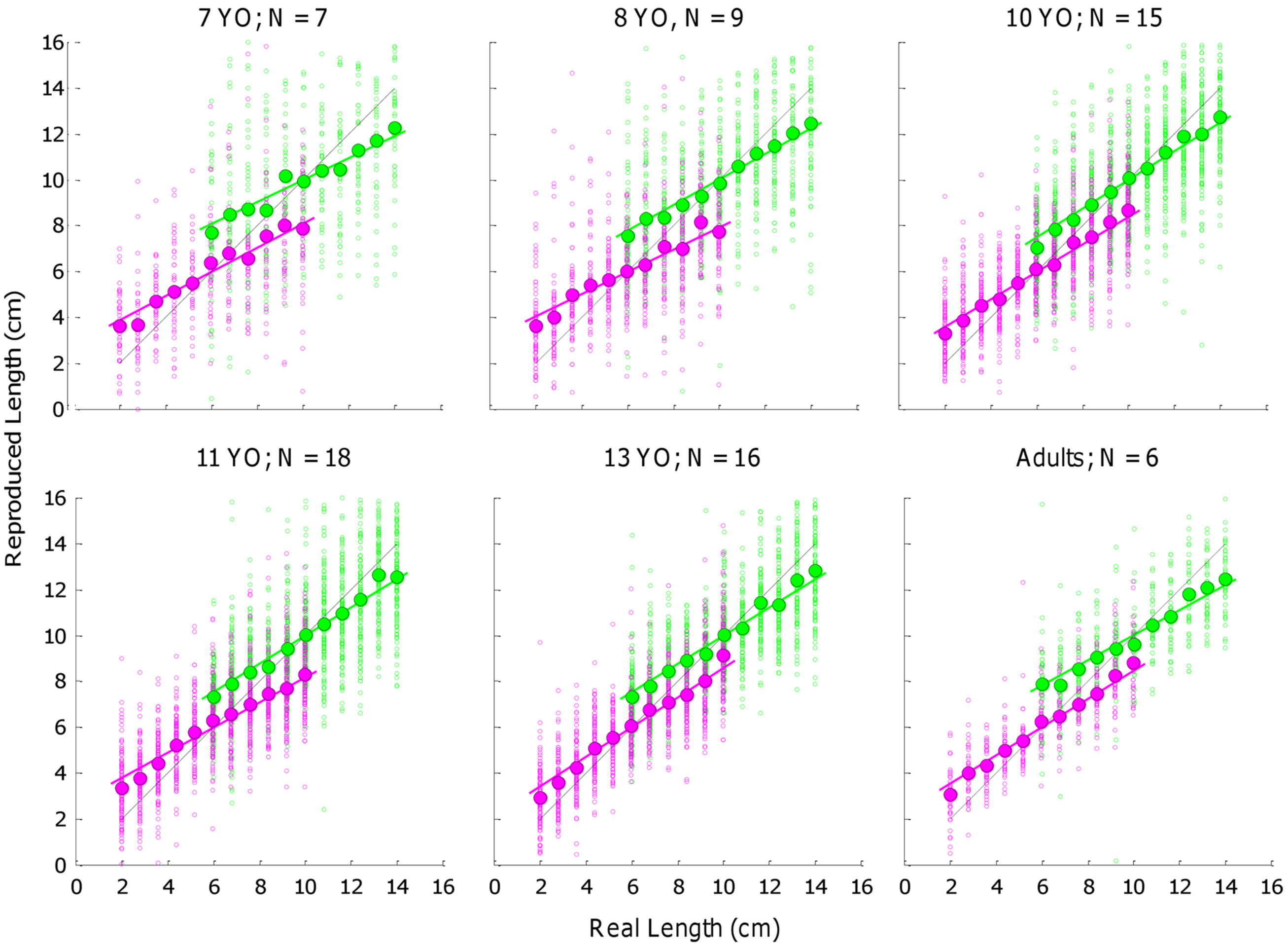


Fig. 3 Length estimates plotted as a function of real stimulus length for the different age group tested (different subplots). *Small dots* represent single estimates, while *larger dots* represent population average estimate for the stimulus. *Magenta* refers to results of the Short session, *green* to those for the Long session. *Dashed black lines* indicated the *identity line*; *continuous magenta/green lines* represent linear fits of the data of the Short and Long sessions, respectively (color figure online)

analysis was aimed at deriving which levels of prior variance would be consistent with the data.

Results

Figure 3 shows the length reproduction for all subjects of each different age ranges, separately for the Short (magenta) and Long (green) interval ranges. These graphs clearly show that the phenomenon of central tendency occurs for space perception, at all ages: The average reproduced lengths (larger dots) are not equal to the real stimulus amplitude, as they do not lie on the identity line, but tend toward the average stimulus length for a particular session (6 and 10 cm for the magenta and green data points, respectively). The reproduced amplitudes have shallower slope than the identity line.

An important consequence of regression toward the mean is that the same interval will be reproduced differently depending on which distribution it was drawn from. The green and magenta dots for lengths are not superimposed between 6 and 10 cm: The same length is reproduced as longer when sampled from the long distribution (10 cm, green) compared with the Short distribution (6 cm, magenta). This phenomenon is more clearly depicted in Fig. 4, which shows that the distribution of the reproduction of the 8.4 cm stimulus amplitude depends strongly on the sample range it was drawn from, with shorter mean amplitude for stimulus drawn from the Short session (magenta) and larger mean amplitude for the Long session (green). Interestingly, all tested age groups showed a tendency to regress toward the mean of the stimulus. At all ages, the regression index was significantly different from 0 ($p < 0.01$ in one-sample t tests), and did

Author's personal copy

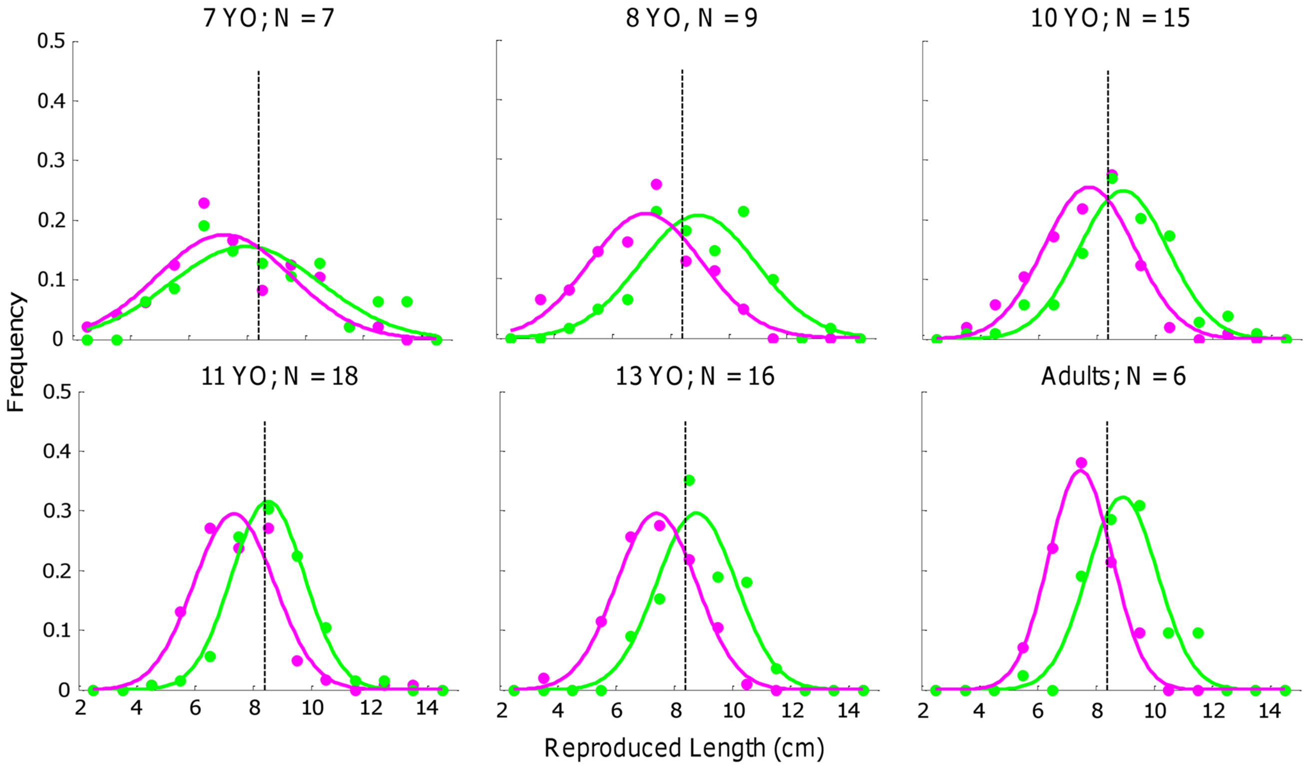


Fig. 4 Distributions of the reproductions of the 8.4 cm length, as a function of stimulus context (Short – *magenta*, Long – *green*) and age group (different subplots) (color figure online)

not vary significantly with age (one-way ANOVA, F(5, 65) = 1.05, $p = 0.397$), slightly decreasing from 7 years on, but always remaining around 0.36 ± 0.06 (SD). Regression indexes were as follows: 0.44 ± 0.23 for the 7-year-olds, 0.40 ± 0.14 for the 8-year-olds, 0.31 ± 0.20 for the 10-year-olds, 0.35 ± 0.10 for the 11-year-olds, 0.29 ± 0.20 for the 13-year-olds and 0.35 ± 0.20 for the adults.

To assess whether subjects modified their tendency to regress toward the mean over the course of the experiment, a split-half analysis of the regression index was conducted on the Long session. Figure 5 clearly shows that children at all ages exhibit a similar significant increase in the regression index over the session, as confirmed by a two-way mixed-model ANOVA (between factor: "age"; within factor: "experimental phase," with two levels: first half, second half. Significant effect of "experimental phase": F(1, 65) = 7.11, $p < 0.01$; not significant effect of "age" or of the interaction: F(5, 65) = 1.14, $p = 0.35$ and F(5, 65) = 0.94, $p = 0.46$, respectively). Although not statistically significant, adults seem to represent an exception to this pattern, as their regression index is similar between the two halves of the task.

Separately, we obtained an estimate of spatial precision using a bisection task, where priors should not influence performance. Subjects reported whether the central flash of a simultaneous triplet was spatially closer to the leftmost or the rightmost one, leading to an estimate of the relative precision of length judgment (Weber fraction). The average Weber fraction decreases (precision increases) substantially with age, with a significant reduction between the fraction

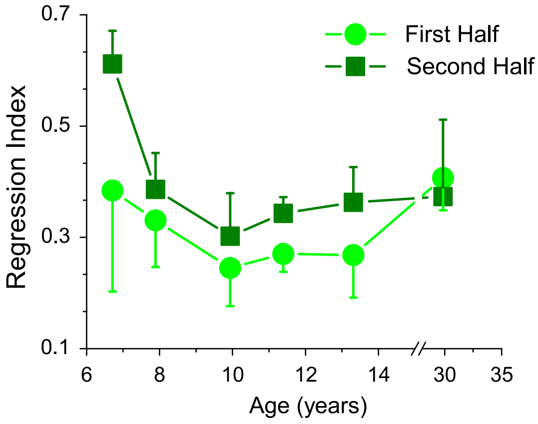


Fig. 5 Average regression index as a function of age in a split-half analysis of the data in the "Long" session. *Different symbols* indicate the two halves of the session: *Circles* represent the regression index computed on the first half, while *squares* refer to the second half of the session. *Error bars* represent standard errors of the mean

Author's personal copy

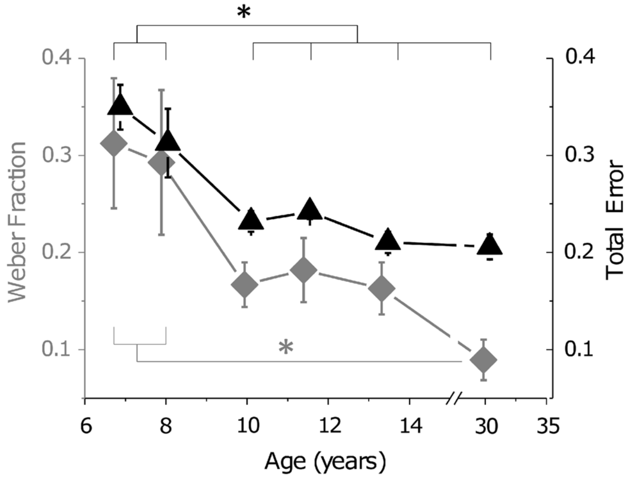


Fig. 6 Average Weber fraction computed from the *length discrimination* task (*gray line* and *symbols*) and average total error computed for the *length reproduction* task (*black line* and *symbols*—see Eq. 4) as a function of age. *Error bars* represent standard errors of the mean. Stars indicate significant difference ($p < 0.05$) in a one-way ANOVA followed by Bonferroni post hoc

measured for the two youngest groups and that assessed in adults [F(5,65) = 3.407, $p < 0.01$ in a one-way ANOVA followed by Bonferroni post hoc test, see gray line and symbols in Fig. 6].

A clear developmental trend was also present in the reproduction task, quantified as the total error (RMSE) normalized by average length. Figure 6 (black line and triangles) shows that the average error decreased significantly with age, dropping from about 35 % at 7 years of age to about 21 % in adulthood. One-way ANOVA followed by a Bonferroni post hoc test showed a significant drop in error between 7 and 8 years, and all the older groups [F(5,65) = 11.379, $p < 0.001$].

Partitioning the error, we could evaluate whether the developmental trend was mainly due to a change in the accuracy of the reproduction (bias) or in precision (coefficient of variation). For all the ages tested, the ratio between the two error components was approximately constant [one-way ANOVA, F(5,65) = 0.85 p = 0.519], with CV being on average 1.577 ± 0.15 (SD) times the bias (see the almost constant angle subtending all symbols in Fig. 7). Therefore, it seems that the perceptual strategy adopted in length reproduction does not substantially change during childhood.

More specifically, we have shown that children show central tendency at all the ages tested (see Figs. 3, 4). Hence, already from 7 years of age, humans base their judgments not only on their current perception, but also on the average of the previously presented stimuli (or the prior). Moreover, the almost constant ratio between inaccuracy (or bias) and variability (CV) in determining total error during childhood suggests that children give proportionally the same weight to the prior as adults do. This would imply that the prior width changes during development, in parallel with the increase in sensory precision. Actually, the relevance given to the stimulus average depends, according to a Bayesian model, on the relative width of the likelihood function (here represented by individual Weber fractions) and the prior (see modeling in the Methods section and Fig. 2 for more details). During development, the Weber fraction diminishes substantially, becoming in adults about one-third of that of 7-year-olds (see Fig. 6).

To evaluate the potential change of prior width, in Fig. 8a, we have plotted the regression indexes for the different age groups measured in the Long condition as a function of their sensory precision (Weber fraction) and in Fig. 8b, the corresponding bias and coefficient of variation. In both graphs, we have added the predictions formulated

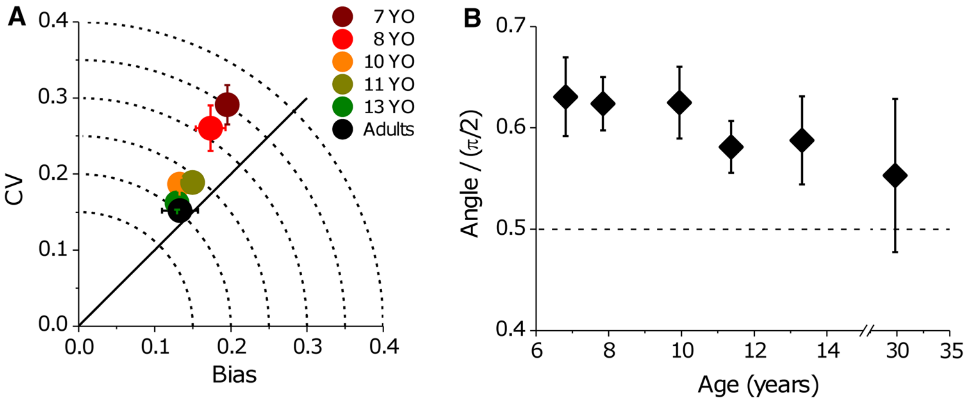


Fig. 7 Partitioning the error. **a** Coefficient of variation of the reproductions (CV, Eq. 3) plotted against bias (difference between average reproduced length and physical sample length, Eq. 2) for the different age groups. The total error (root mean squared error, Eq. 4) is given by the distance from the origin. *Error bars* represent group standard errors of the mean. **b** Average *polar angle* between bias and coefficient of variation of the reproductions as a function of age. *Error bars* represent standard errors of the mean

Author's personal copy

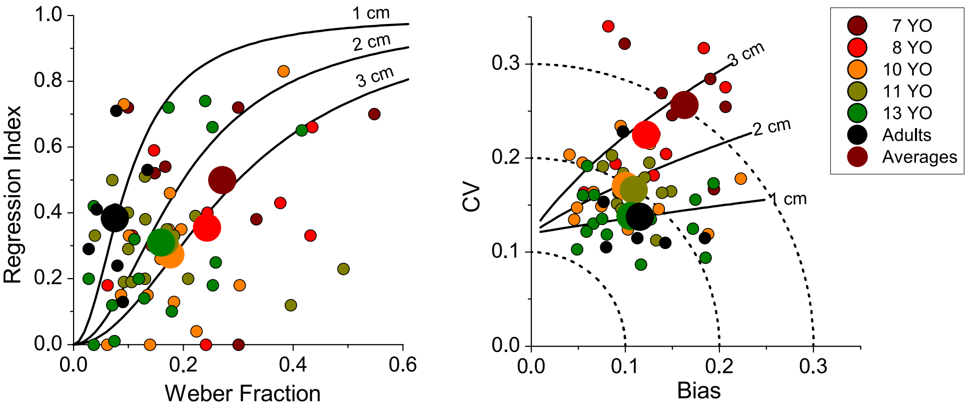


Fig. 8 Model simulations. **a** Regression index plotted against Weber fraction and **b** coefficient of variation plotted against bias for the data in the Long session. *Small circles* are individual subjects; *large circles* are group averages for different ages (same *color code* as Fig. 7). The *black curves* show the Bayesian model predictions for a prior of fixed width of 1, 2 and 3 and assuming a fixed additive motor noise approximated as Gaussian, with 1.2 cm width (selected to fit adults average data). Each *curve* was created by varying sensory Weber fraction from 0.01 to 0.6

for different prior widths (black lines). Our results seem compatible with the hypothesis of a prior width undergoing a development similar to that observed for sensory precision, yielding an almost constant proportion between prior and likelihood amplitude. Indeed, looking at Fig. 8, it emerges that models with a decreasing prior width (see lines corresponding to priors from 3 to 1 cm) progressively approximate better the regression indexes (Fig. 8a) and the error components (Fig. 8b) of the different age groups (large filled circles). Therefore, already at 7 years of age, the relative weighting of sensory information and prior is similar to that measured in adults.

Discussion

Our experiments show that the central tendency mechanism is already active during childhood, at least from 7 years of age on. Indeed, participants at all the ages tested tended to reproduce the presented length differently as a function of the previous stimuli history: The same stimulus length was reproduced as shorter, when on average stimuli were short and as longer, when average stimuli were long. We have provided evidence that this strategy, which reduces the variability of the responses at the expenses of accuracy and reflects a Bayesian optimization principle (Jazayeri and Shadlen 2010; Cicchini et al. 2012), is adopted since early childhood.

It is well known that children can encode the statistical properties of the world. Evidence exists that even young infants show statistical learning, the ability to extract the statistical structure of the sensory inputs (Bulf et al. 2011). For instance, 5-month-old infants are sensitive to statistical information of colors and shapes (Kirkham et al. 2002), 8-month-old infants can compute statistics in spatiotemporal visual sequences (Kirkham et al. 2007) and exploit the statistical relations to segment continuous speech into words (Saffran et al. 1996). Slightly older infants (9-month-olds) encode the statistical structure of scenes constituted by multiple elements, being sensitive not only to the frequency of co-occurrence of different objects, but also to the predictability between elements, i.e., to their conditional probability relations (Fiser and Aslin 2002). Hence, even during the first year of life, infants can take advantage of the regularities in the environment to help their understanding of complex stimuli. This ability can be traced back, at least partially, even to newborns, as neonates can learn the statistical structure of a visual sequence of discrete shapes (Bulf et al. 2011) or of a linguistic stimulus (Teinonen et al. 2009).

However, children do not always exploit statistics as adults do. During childhood, children show a significant physical and sensory development, which requires continuous motor and perceptual re-calibration. For example, the association between perceptions and the true object characteristics needs to be learned and recalibrated as various aspects of the body (such as limb size and inter-ocular distances) change (Gori et al. 2008). In this phase, the veridicality of an estimate plays a fundamental role, to allow for the development of an accurate representation of the real world. Hence, during such a calibration process, the precision of the estimate (the variability associated with the perception) could be considered less relevant, inducing children to base their perceptual judgments on their sensory input without relying on statistically optimal integration strategies. This is what happens, for instance, with sensory cue integration. In adults, cues from different modalities (e.g., vision and touch or sound, Ernst and Banks 2002;

Author's personal copy

Alais and Burr 2004) or even from within the same modality (e.g., visual cues as stereo and texture, Jacobs 1999; Knill and Saunders 2003) are integrated in an optimal fashion, well modeled in a Bayesian framework. This strategy privileges, among the multiple sources of information, the more precise sensory cue. However, children use multisensory information in a substantially different way from adults (Gori et al. 2008; Nardini et al. 2008). This has been, for instance, demonstrated for visuo-haptic size perception, where until 8–10 years of age children's estimate is not based on the more precise (i.e., less variable) sense, but seems to reflect a haptic dominance (Gori et al. 2008). Children would therefore rely on the more accurate signal (haptic), which acts as a calibrator for the other senses. This interpretation has recently found further support in studies conducted on patient populations, where the absence of the "calibrator modality" during childhood (e.g., of the haptic sense for size perception) has determined impairments in estimates performed also with other modalities (e.g., visual size estimates) (Gori et al. 2010, 2012b).

If a similar need for calibration had influenced the integration between current sensory inputs and prior knowledge, we might not have expected the consistent regression to the mean that we have actually measured. On the contrary, the development of the incorporation of prior knowledge in spatial judgments seems to occur early in childhood, being active already in 7-year-olds.

A question that can be raised is therefore why the development of central tendency occurs so early. A possible answer could be found in the relevance of this mechanism—and more in general of a process incorporating prior knowledge in perception—for learning. The use of previous knowledge (or priors) is a fundamental tool to deal with the uncertainty associated with sensation, where the variability of the signal is often due to noise (neural or sensory) rather than to a real variation in the physical world (Knill and Pouget 2004). The inability to build and use prior knowledge to filter out such noise could hinder the ability to learn and generalize sensory information, preventing the child to derive the "general trend" of the signal. A similar phenomenon, a low weighting given to the priors (or hypo-priors), has been recently suggested as a possible cause for several non-social symptoms of autism (Pellicano and Burr 2012). Even more importantly, the use of priors (or assumptions) becomes essential for perception in the case of ambiguous sensory inputs, where additional information is required to disambiguate incoming information. For instance, the univocal recovery of three-dimensional information from a two-dimensional retinal image is an ill-posed problem, as the same object when projected in two dimensions can yield to different images. It is just with additional assumptions (or priors) that images can be disambiguated. A common example is that of the "light-from-above" prior, where the unconscious assumption that objects are generally illuminated from above, makes us perceive shaded two-dimensional images as three-dimensional objects, resolving the convex/concave ambiguity (Rittenhouse 1786; Brewster 1826). Similarly, we would not be able to catch a falling object, if we did not exploit an internal prior about the effects of gravitational force on its motion to anticipate object acceleration, which is not well perceived by our visual system (McIntyre et al. 2001). Therefore, the ability to encode the statistics of the environment and incorporate such "previous knowledge" for perception and learning could represent an essential trait to allow for the successive development of more complex perceptual and action skills.

How this previous knowledge is generated and how it changes during life are, however, still two open questions. Our findings address these interrogatives, suggesting that their formation is characterized by two different time courses. On the one hand, a few minutes of training allows for the development of a new prior (i.e., the average amplitude of the presented stimuli), which then influences the immediate estimate of the length of other stimuli. The possibility to build (and use) new priors in 5–10 min is compatible with what has been recently observed for time perception in adults (Jazayeri and Shadlen 2010; Cicchini et al. 2012) and also with the finding that also existing priors (as the "light-from-above" one) can be altered by short training (Adams et al. 2004). On the other hand, the prior width, on which the relative weight of the prior in the total percept depends (if we assume that all observers follow a reliability-based optimal weighting strategy, see Bayesian modeling in the "Methods" section and Fig. 2), undergoes a change with age. In particular, the estimates measured for young children are compatible with a prior width much larger than that fitting adult data. Therefore, the absolute width of priors can change during development, yielding to an almost constant relative weighting as sensory precision improves.

It is important to note that in the present study, we cannot prove that children are using an optimal weighting strategy, because the optimal weights for prior and sensory inputs cannot be estimated as the variance of the prior distribution is unknown. The same data could alternatively be explained by a sub-optimal behavior (as, for instance, by a misweighting, in which similar weights are given to prior and sensory information independently of their reliability). Our analysis was aimed at deriving which levels of prior variance would be consistent with the data, if observers used an optimal weighting strategy. Keeping this assumption in mind, the results show that with age, sensory precision increases (Weber fraction decreases), while relative reliance on the prior stays constant, a result consistent with performance at all ages being based on an optimally weighted average of sensory and prior information. This phenomenon can

Author's personal copy

be better described by two components: The development of the underlying statistical learning mechanism and the quality of sensory information available to build a prior. Indeed, perceptual precision in spatial perception significantly improves with age, as shown by the reduction of the Weber fraction measured in the length discrimination task (see Fig. 6). Therefore, noisier sensory information would lead to a wider prior in younger children, while more precise perception would allow the construction of a tight prior. Thus, a potential interpretation of the observed developmental trend is that the major limitation on the quality of the prior that can be acquired is the quality of the sensory information available. Hence, a possibility is that children are about as good at the statistical learning component of this task as adults, but are just receiving noisier information with which to build their prior on each trial.

A similar trend of change in prior width as a function of age has been recently pointed out by Stone (2011) for children between 4 and 11 years of age in the context of the "light-from-above" prior. As previously mentioned, adults tend to assume lighting comes from above to judge the convexity/concavity of a shaded image. This prior is apparently applied also by infants (at least for artificial stimuli, Granrud et al. 1985) and young children, from 3 to 4 years of age (Yonas et al. 1979; Stone and Pascalis 2010). However, the propensity to assume that lights come from above (that is, the prior weight) gradually changes with age and shifts toward an adultlike prior value as children grow older (e.g., between 4 and 12 years, as reported by Thomas et al. (2010) for "polo-mint" stimuli). Our findings suggest that a similar developmental trend does not regard exclusively the environmental priors, which could derive from our lifelong experience with the light coming from above us, or of the gravitational force field acting on our body. On the contrary, also the priors formed on a brief time scale on the basis of a specific stimulus history apparently undergo a progressive change with age. This way, a balance is maintained between sensory inflow and prior information, by complementing a noisier perception in younger children with a weaker prior and a more precise sensory input in adulthood with a proportionally stronger one. Previous research in adults has shown that the adoption of this strategy is optimal in Bayesian terms, as it allows systems to compensate for reduced sensory resolution, minimizing total error in a temporal reproduction (Jazayeri and Shadlen 2010; Cicchini et al. 2012). This strategy is apparently successful in holding the sensory noise at bay for length reproduction, even during development, since children's performance reaches adult levels in our task by 10 years of age in terms of total error (see Fig. 6, black line).

An interesting question could be to investigate the time course of the statistical learning. As mentioned before, children are very quick to acquire priors based on a few stimuli (as the lengths used in our experiment), but fairly slow to acquire priors based on stimuli that are constantly experienced, such as the light-from-above. This difference suggests the existence of separate learning mechanisms, one characterized by a rapid acquisition and a rapid forgetting of the information, and the other producing a long-lasting internalization of the prior, but requiring a much longer exposure to develop. Even considering the fastest phenomenon, how many stimuli should an individual perceive to be able to build a prior distribution and use it in his subsequent estimations? Our results (see Fig. 5) indicate that for children, the prior can change already over the course of a short task (77 trials), acquiring progressively more relevance between the first and the second half of experiment. Interestingly, we might speculate that such learning process is even faster for adults, who apparently learned the prior already early at the beginning of the experiment and did not modify it substantially between the first and the second half of the trials. Recent findings confirm that this process could be very fast, at least for the perception of number, with a weighted average just with the previous trial being enough to bias perception, even in children (Cicchini et al. 2014). It would be interesting to see if this also explains the effects of spatial context dependency in children.

A possible confound in this study could be represented by the fact that we did not account for the change in motor error that is potentially associated with development. However, an increased motor noise would be expected to raise reproduction variability, but would not substantially influence the central tendency (the bias and the regression index). In graphical terms, a larger motor noise would produce essentially an upward shift of the model lines in Fig. 8b (which are currently computed with a fixed motor noise approximated as a Gaussian with standard deviation of 1.2 cm), with no (or minor) impact on the model lines in Fig. 8a. Therefore, although the definition of a complete model of the reproductions would require a direct measure of motor noise, the conclusions that we draw from the current experiments do not depend on its exact knowledge and maintain their validity also assuming its decrease with increasing age.

As Helmholtz and Gregory argued long ago (Helmholtz 1866; Gregory 1980), perception is a process of unconscious inference, where prior knowledge has a fundamental influence on what we actually see. Here, we have shown that one aspect of this mechanism, central tendency (Hollingworth 1910), is already active in young children. The performance of seven-year-olds is consistent with a strategy that exploits an estimate of the stimulus history to bias perceptual judgments, in order to increase the reliability of the responses and reduce overall error. The early occurrence of influence of previous knowledge on the perception of even very simple stimuli (such as length or duration)

Author's personal copy

emphasizes how this mechanism is widespread already in the developing child. It is therefore a key issue for the future to understand how this mechanism is actually implemented in the brain (see Fiser et al. 2010 for a review), and whether its dysfunction can account for the complex symptoms associated with neuropsychiatric conditions, as some authors suggest—e.g., for autism (Pellicano and Burr 2012) or schizophrenia (Fletcher and Frith 2009).

Acknowledgments This study was supported by the ABBI (FP7-ICT-611452) and the CODEFROR (PIRSES-2013-612555) European projects.

References

Adams WJ, Graf EW, Ernst MO (2004) Experience can change the 'light-from-above' prior. Nat Neurosci 7:1057–1058

Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. Curr Biol 14:257–262

Atkinson J (2000) The developing visual brain. Oxford University Press, New York

Brainard DH (1997) The psychophysics toolbox. Spat Vis 10:433–436

Brewster D (1826) On the optical illusion of the conversion of cameos into intaglios, and of intaglios into cameos, with an account of other analogous phenomena. Edinb J Sci 4:99–108

Bulf H, Johnson SP, Valenza E (2011) Visual statistical learning in the newborn infant. Cognition 121:127–132

Burr D, Gori M (2011) Multisensory integration develops late in humans. In: Murray MM, Wallace MT (eds) Frontiers in the Neural Bases of Multisensory Processes. Taylor & Francis Group, USA

Cicchini GM, Arrighi R, Cecchetti L, Giusti M, Burr DC (2012) Optimal encoding of interval timing in expert percussionists. J Neurosci 32:1056–1060

Cicchini GM, Anobile G, Burr DC (2014) Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. PNAS (in press)

Efron B, Tibshirani RJ (1993) An Introduction to the Bootstrap. Chapman and Hall, New York

Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415:429–433

Fiser J, Aslin RN (2002) Statistical learning of new visual feature combinations by infants. Proc Natl Acad Sci USA 99:15822–15826

Fiser J, Berkes P, Orban G, Lengyel M (2010) Statistically optimal perception and learning: from behavior to neural representations. Trends Cogn Sci 14:119–130

Fletcher PC, Frith CD (2009) Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. Nat Rev Neurosci 10:48–58

Gori M, Del Viva M, Sandini G, Burr DC (2008) Young children do not integrate visual and haptic form information. Curr Biol 18:694–698

Gori M, Sandini G, Martinoli C, Burr D (2010) Poor haptic orientation discrimination in nonsighted children may reflect disruption of cross-sensory calibration. Curr Biol 20:223–225

Gori M, Squeri V, Sciutti A, Masia L, Sandini G, Konczak J (2012a) Motor commands in children interfere with their haptic perception of objects. Exp Brain Res 223:149–157

Gori M, Tinelli F, Sandini G, Cioni G, Burr D (2012b) Impaired visual size-discrimination in children with movement disorders. Neuropsychologia 50:1838–1843

Granrud CE, Yonas A, Opland EA (1985) Infants' sensitivity to the depth cue of shading. Percept Psychophys 37:415–419

Gregory RL (1980) Perceptions as hypotheses. Philosophical transactions of the royal society of London. Ser B Biol Sci 290:181–197

Helmholtz H (1866) Handbuch der physiologischen Optik. Voss, L., Leipzig

Hollingworth HL (1910) The central tendency of judgements. J Philos Psychol Sci Methods 7:461–469

Jacobs RA (1999) Optimal integration of texture and motion cues to depth. Vision Res 39:3621–3629

Jazayeri M, Shadlen MN (2010) Temporal context calibrates interval timing. Nat Neurosci 13:1020–1026

Kersten D, Yuille A (2003) Bayesian models of object perception. Curr Opin Neurobiol 13:150–158

Kirkham NZ, Slemmer JA, Johnson SP (2002) Visual statistical learning in infancy: evidence for a domain general learning mechanism. Cognition 83:B35–B42

Kirkham NZ, Slemmer JA, Richardson DC, Johnson SP (2007) Location, location, location: development of spatiotemporal sequence learning in infancy. Child Dev 78:1559–1571

Kleiner M, Brainard D, Pelli D (2007) What's new in Psychtoolbox-3? Perception 36

Knill DC, Pouget A (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci 27:712–719

Knill DC, Saunders JA (2003) Do humans optimally integrate stereo and texture information for judgments of surface slant? Vision Res 43:2539–2558

McIntyre J, Zago M, Berthoz A, Lacquaniti F (2001) Does the brain model Newton's laws? Nat Neurosci 4:693–694

Mennemeier M, Pierce CA, Chatterjee A, Anderson B, Jewell G, Dowler R, Woods AJ, Glenn T, Mark VW (2005) Biases in attentional orientation and magnitude estimation explain crossover: neglect is a disorder of both. J Cogn Neurosci 17:1194–1211

Morrongiello BA, Humphrey GK, Timney B, Choi J, Rocca PT (1994) Tactual object exploration and recognition in blind and sighted children. Perception 23:833–848

Nardini M, Jones P, Bedford R, Braddick O (2008) Development of cue integration in human navigation. Curr Biol 18:689–693

Pellicano E, Burr D (2012) When the world becomes 'too real': a Bayesian explanation of autistic perception. Trends Cogn Sci 16:504–510

Rentschler I, Juttner M, Osman E, Muller A, Caelli T (2004) Development of configural 3D object recognition. Behav Brain Res 149:107–111

Rittenhouse D (1786) Explanation of an optical deception. Trans Am Philos Soc 2:37–42

Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. Science 274:1926–1928

Stone JV (2011) Footprints sticking out of the sand. Part 2: children's Bayesian priors for shape and lighting direction. Perception 40:175–190

Stone JV, Pascalis O (2010) Footprints sticking out of the sand. Part 1: children's perception of naturalistic and embossed symbol stimuli. Perception 39:1254–1260

Streri A (2003) Cross-modal recognition of shape from hand to eyes in human newborns. Somatosens Mot Res 20:13–18

Teinonen T, Fellman V, Naatanen R, Alku P, Huotilainen M (2009) Statistical language learning in neonates revealed by event-related brain potentials. BMC Neurosci 10:21

Thomas R, Nardini M, Mareschal D (2010) Interactions between "light-from-above" and convexity priors in visual development. J Vis 10:6

Watson AB, Pelli DG (1983) QUEST: a Bayesian adaptive psychometric method. Percept Psychophys 33:113–120

Weiss Y, Simoncelli EP, Adelson EH (2002) Motion illusions as optimal percepts. Nat Neurosci 5:598–604

Yonas A, Kuskowski M, Sternfels S (1979) The role of frames of reference in the development of responsiveness to shading information. Child Dev 50:495–500