View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Nottingham Trent Institutional Repository (IRep)

Running head: Context and timing in children and adults

The dynamic effect of context on interval timing in children and adults

Quentin Hallez[1], Atser Damsma[2], Darren Rhodes[3,4], Hedderik Van Rijn[2], Sylvie Droit-Volet[1]

[1] University Clermont Auvergne, CNRS, Clermont-Ferrand, France

[2] University of Groningen, Groningen, Netherlands

[3] University of Sussex, Brighton, United Kingdom

[4] Nottingham Trent University, Nottingham, United Kingdom

Corresponding authors: Quentin Hallez, Université Clermont Auvergne, Laboratoire de psychologie sociale et cognitive, CNRS, UMR 6024, 34 avenue carnot, 63000 Clermont-Ferrand. E-mail. quentin.hallez@gmail.com.

Abstract: Human reproductions of time intervals are often biased towards previously perceived durations, resulting in a central tendency effect. The aim of the current study was to compare this effect of temporal context on time reproductions within children and adults. Children aged from 5 to 7 years, as well as adults, performed a ready-set-go reproduction task with a short and a long duration distribution. A central tendency effect was observed both in children and adults, with no age-difference in the effect of global context on temporal performance. However, the analysis of the effect of local context (trial-by-trial) indicated that younger children relied more on the duration (objective duration) presented in the most recent trial than adults. In addition, statistical analyses of the influence on temporal performance of recently reproduced durations by subjects (subjective duration) revealed that temporal reproductions in adults were influenced by performance drifts, i.e., their evaluation of their temporal error, while children simply relied on the value of reproduced durations on the recent trials. We argue that the central tendency effect was larger in young children due to their noisier internal representation of durations: A noisy system led participants to base their estimation on experienced duration rather than on the evaluation of their judgment.

Keywords: Development, Temporal context, Reproduction, Bayesian Timing, Central Tendency, Decision-Making

1. Introduction

We live in a dynamic world with a plurality of temporal events and some of them that might fluctuate in their temporal properties, going faster or slower than usual. Given that time is a fundamental dimension of perception, action and cognition, we can assume that humans continuously adjust their behaviour to these changing temporal properties of our physical environment (Di Luca & Rhodes, 2016; Rhodes, 2018). The acquisition of the duration associated to an event therefore depends on the temporal context of learning (Rattat & Tartas, 2017). This paper tests the degree to which our prior knowledge about the temporal properties of the world is learnt and used at different developmental ages.

It is well documented in studies with human adults that temporal context influences the estimation of different magnitudes, including temporal rhythm or duration (Adams & Mamassian 2004; Battaglia, Jacobs & Aslin, 2003; ; Damsma, van der Mijn & van Rijn, 2018; Ernst & Banks, 2002; Jazayeri & Shadlen, 2010; Körding, Beierholm, Ma, Quartz, Tenenbaum & Shams, 2007; McAuley & Jones, 2003, McAuley, Jones, Holub, Johnston & Miller, 2006; Mamassian, Landy & Maloney, 2002; Miyazaki, Nozaki & Nakajima, 2005; Petzschner, Maier & Glasauer, 2012; Shi & Burr, 2016; Stocker & Simoncelli, 2006; Verstynen & Sabes, 2011). This phenomenon is illustrated by the central tendency effect described by Hollingworth (1910), and known in the psychology of time as Vierordt's (1868) law. According to Vierordt's law, in a task in which a range of time intervals have to be reproduced, participants tend to overestimate the shortest durations and underestimate the longest durations (Lejeune & Wearden, 2009). This bias in time estimates demonstrates that the judgment of durations is not absolute, but relative to the centre of the distribution of tested durations. The judgment of a given duration therefore depends on the previous encountered durations.

According to the Bayesian theory of perceptual inference for time, the currently perceived interval (the likelihood) is weighted with previous experience (the prior) to come to a subjective estimation of duration (the posterior). So, in a temporal task with a sequence of trials, there would be an "online prior" where the prior is updated on a trial-by-trial basis, with a greater influence on the current estimate of more recent trials (Dyjas, Bausenhart & Ulrich, 2012; Di Luca & Rhodes, 2016; Lapid, Ulrich & Rammsayer, 2008; Taatgen & Van Rijn, 2011; van Rijn, 2016). In addition, the Bayesian view predicts that the noisier the time estimates are, the more participants will rely on prior knowledge. As explained by Jazayeri and Shadlen (2010, p. 1020), "the brain takes into account knowledge of temporal uncertainty and adapts its time keeping mechanisms to temporal statistics in the environment". Indeed, given that the standard deviation of temporal judgment increases with the length of durations to be estimated, as indicated the scalar property of timing (for a review see Wearden, 2016), it has been found that the central tendency effect is stronger for longer stimulus durations (Cicchini, Arrighi, Cecchetti & Burr, 2012; Jazayeri & Shadlen, 2010).

The scalar variability of timing has been verified in young children in different tasks (for recent reviews see Droit-Volet, 2013, 2016; Coull & Droit-Volet, 2018). In addition, the variability in estimates has been systematically shown to be higher in young children than in adults. We can therefore assume that the uncertainty in time judgments is higher in younger children, and as such, they might rely on prior experience to a greater extent than adults do. The few developmental studies on temporal reproduction showed a stronger temporal bias in children, with a higher over- and underestimation of short and long durations, respectively (Crowder & Hohle, 1970; Droit-Volet, Wearden & Zélanti, 2015; Szelag, Kowalska, Rymarczyk & Pöppel, 2002). This typical temporal bias has been explained by the motor component of this task (Droit-Volet, 2010). The higher overestimation of short durations in young children compared to adults would be due to their motor responses that took more time to complete, while the higher underestimation of long durations might be due to their motor impulsivity. In line with these findings, some authors have warned against using this temporal task in young children (Droit-Volet, 2010; Indraccolo, Spence, Vatakis & Harrar, 2016). However, although the contribution of motor action in age-related differences in temporal reproduction cannot be excluded, we can also assume a stronger effect of prior knowledge on temporal reproduction in young children than in adults.

A recent study using the temporal reproduction task has been conducted in autistic and typically developed children aged from 6 to 14 years (Karaminis et al., 2016). The results replicated the central tendency effect in all age groups, with a stronger effect for younger participants. In addition, Bayesian modelling of the data suggested a higher reliance on the prior in young children than in adults. The autistic children showed a lower sensitivity to time, but did not rely more on prior knowledge than age-matched typical children to compensate for their temporal error. However, as reported the authors, unexpectedly, the context dependent effect was not consistent across age groups, being absent in children older than 10 years and adults (p. 3). This is likely due to the fact that younger children underestimated all durations, thereby reducing the context effect to which they may be subject (Hallez & Droit-Volet, 2017; Karaminis et al., 2016).

The aim of the present study was to replicate and extend these results on the effect of temporal context on temporal reproduction performance in children as young as 5 years old. Indeed, the originality of our study lays on the examination of the influence of temporal performance in children and adults. The *global* context (i.e., the range of presented intervals) was not the only focus however, as we also investigated the *local* context (i.e., the direct effect of recent trials), a distinction that has not yet been investigated from a developmental perspective. In the present study, children aged 5, 6 and 7 years, as well as adults, performed a "ready-set-go" reproduction task in which we manipulated the temporal context by using two different ranges of durations: a short and a longer range. To assess the effect of this global context manipulation, one duration in the two temporal ranges overlapped. We hypothesized an effect of temporal context on temporal performance for both children and adults, with the overlapping duration judged longer in the long than in the short context condition. In addition, because of the lower temporal sensitivity in young children, we expected that the effect of recent prior trials would be higher in children than in adults.

2. Methods

2.1. Participants

A total of 24 five-year-olds (11 females), 31 six-year-olds (16 females), and 25 seven-year-olds (10 females) and 33 adults (27 females, mean age = 20.43, *SD* = 3.94) took part in this experiment. Children were recruited from different nursery and primary schools, whereas adults were Psychology students of the University Clermont Auvergne, all located in the municipality of Clermont-Ferrand, France. Children’s parents as well as adult participants signed written informed consent for their participation in this experiment, which was carried out according to the principles of 1964 Helsinki’s declaration and approved by the academy committee of the French National Education Ministry, and the ethics committee of research IRB-UCA, according to ethical standards of the French law.

2.2. Apparatus and stimuli

In a quiet room, participants were seated in front of a cathode screen on which all stimuli were presented. The screen was linked to a MSI Apach Pro computer that launched all experimental events and recorded responses using Psychtoolbox-3 (Brainard, 1997; Kleiner et al., 2007) in Matlab.

During an entire experimental block, a 0.8° fixation cross was presented at the centre of the screen (Figure 1). In each trial, a *warning*, *ready* and *set* stimulus were presented. The warning stimulus consisted of a 2.0° diameter black circle with the label 'ready', and appeared on the left of the fixation cross at a random distance between 4.0° and 8.1°. The ready and set stimuli consisted of a white 2.0° diameter circle. The ready circle was presented on the right of the fixation cross at a random distance between 4.0° and 8.1°. The set circle was always located 4.8° above the fixation cross.

2.3. Procedure

All participants performed a ready-set-go reproduction task in two temporal contexts: one with short durations and the other with long durations. The presentation order of this context condition was counterbalanced across participants. The fulfilment of each of the two conditions was done on two distinct days. The 0.9 s interval duration was presented in each contextual condition, in order to examine whether the temporal reproduction of this target duration was affected by the temporal context. In the “short” context condition, the interval duration were 0.5, 0.6, 0.7, 0.8 and 0.9 s, and the “long” context condition 0.9, 1, 1.1, 1.2, 1.3 s. In each condition, the participants were given 4 blocks of 20 trials (a total of 80 trials), that is 8 trials per interval duration. The presentation order of the interval durations was random. Participants were given a demonstration before each temporal condition composed of 10 trials (5 demonstrations and 5 practice trials), in which each duration of the context conditions was presented twice.

Each trial started with a 1 s fixation cross (Figure 1). Then, the black *warning* circle was presented to indicate that a new trial had started. This circle stayed on the screen during the rest of the trial until the participant made a response. After a random interval between .25 and .85 s, the white *ready* circle was presented for 0.1 s, marking the onset of the interval. Next, the offset of the interval was indicated by the presentation of the white *set* circle for 0.1 s. The task of the participants was to immediately reproduce this interval after the presentation of the set circle by pressing spacebar to indicate the offset.

Insert Figure 1 about here

2.4. Data analysis

A complete overview of the analyses and results can be found at osf.io/k3znf. For data analysis, we excluded reproductions lower than 0.1 s and higher than 2.0 s, leading to the exclusion of 6.0% of the total data (12.4, 8.3, 4.5 and 0.4% of the trials for the 5-, 6-, 7-year-olds and adults, respectively). We modeled the data using Linear Mixed Models (LMMs) using the *lme4* package in R (Bates, Maechler, Bolker & Walker, 2014). To test the overall effect of fixed factors, we did model comparisons using likelihood ratio tests. If a fixed factor improved the model fit, it was included. To make the interpretation of the effect of objective duration more straightforward, we centered this continuous factor by subtracting the middle interval (i.e., 0.9 s) from all values. *Subject* was always included as a random intercept term. Next, we sequentially added random slopes for the significant fixed factors to the best model and compared the more complex model with the simpler model using a likelihood ratio test. Random slope terms were included if they improved the model. Post-hoc multiple comparisons were computed using the *glht* function from the *multcomp* package (Hothorn et al., 2013) and the *lsmeans* function from the *lsmeans* package in R (Lenth, 2016).

To quantify the evidence in favor of the null hypotheses (i.e. there is no effect of the particular fixed factor), we calculated Bayes factors using the *lmBF* function from the *BayesFactor* package in R (Morey, Rouder & Jamil, 2014). We will denote the evidence for the null hypothesis (H_0) over the alternative hypothesis (H_1) as BF_{01}.

3. Results

3.1. Mean in temporal reproduction

Figure 2A shows the mean reproduction of interval durations for the different age groups. As can be seen in Figure 2, the children overall showed a smaller slope and a larger underestimation of longer intervals. We modelled the data starting with an LMM predicting reproduction, with subject as a random intercept term. We found adding centered objective duration improved the model fit ($\chi^2(1) = 558.43$, $p < 0.001$, $BF_{01} < 0.01$), showing that overall there was a positive, linear increase of reproductions with objective duration ($\beta = 0.25$, $t = 23.89$, $p < 0.001$). However, adding age group and the interaction between age group and objective duration to the model also improved the model fit ($\chi^2(3) = 38.72$, $p < 0.001$, $BF_{01} < 0.01$ and $\chi^2(3) = 1110.90$, $p < 0.001$, $BF_{01} < 0.01$, respectively), indicating that there was a difference between the age groups in the intercept and slope of the reproductions. Post-hoc multiple comparison showed that the intercept (i.e., the reproduction of 0.9 s estimated by the model) was higher for the adults compared to the 6-year-olds and 7-year-olds (*p*s < 0.001). In addition, the intercept of the 5-year-olds was higher than that of the 7-year-olds ($p < 0.001$). There were no other intercept differences between the age groups (*p*s > 0.078). A second post-hoc test showed that the slope was larger for the adults compared to the three children groups (*p*s < 0.001), but there were no differences between the children groups (*p*s > 0.495).

Insert Figure 2 about here

3.2. Variance in temporal reproduction

We used the coefficient of variation (CV) as measure of the variability in temporal reproductions. To this end, we calculated the CV per subject for each objective duration, as the standard deviation of the average reproduction divided by the average reproduction. Figure 2B shows the average CV per age group. An LMM predicting CV showed that age group improved the fit significantly ($\chi^2(3) = 96.76$, $p < 0.001$, $BF_{01} < 0.01$). A post-hoc Tukey's HSD test showed that relative to all children groups, the adults had a smaller CV (*p*s < 0.001). In addition, the 7-year-olds had a significantly smaller CV than the 5 year olds ($\beta = -0.08$, $z = -3.48$, $p = 0.003$). All other comparisons were non-significant (*p*s > 0.110). Thus, in summary, our results indicate that the CV decreased with age.

3.3. Global context effect

To test whether temporal reproductions were influenced by the global context manipulation, we compared the reproductions of the short and the long context for the overlapping duration (i.e., 0.9 s). Figure 3 shows the average difference between the short and the long context at this interval duration for the different age groups. We found that, overall, the temporal context predicted the reproductions of the overlapping interval significantly ($\chi^2(1) = 31.42$, $p < 0.001$, $BF_{01} < 0.01$). Adding age group to the model improved the fit ($\chi^2(3) = 33.66$, $p < 0.001$, $BF_{01} < 0.01$), indicating the reproduction differed significantly between age groups. Post-hoc comparisons showed that the reproductions at the overlapping interval were significantly longer for the long context compared to the short context for the 5-year-olds ($\beta = 0.09$, $t = 2.15$, $p = 0.033$) and the adults ($\beta = 0.07$, $t = 1.99$, $p = 0.049$). There was no significant difference for the 6- and the 7-year-olds (*p*s > 0.130). Crucially, however, model comparison showed that the effect of context did not differ significantly between age groups ($\chi^2(1) = 4.26$, $p = 0.235$, $BF_{01} = 67.15$).

Insert Figure 3 about here

3.4. Local context effects

3.4.1. Objective previous durations

To quantify the influence of previous presented durations on the current reproduction, we started with the model established previously, including reproduction as the dependent variable and objective duration, age group and context as fixed factors. In addition, the interaction between age group and context and age group and objective duration were included. To this model, we sequentially added objective previous durations (N-1, N-2, N-3, etc.). We found that N-1 and N-2 had a significant influence on the current reproduction ($\chi^2(1) = 37.15$, $p < 0.001$, $BF_{01} < 0.01$ and $\chi^2(1) = 4.56$, $p = 0.033$, $BF_{01} = 0.76$ respectively). However, N-3 did not improve the model fit ($\chi^2(1) = 0.28$, $p = 0.594$, $BF_{01} = 7.54$), so no previous durations beyond N-2 were included in the model.

Figure 4A shows the weight of the previous four objective trials on the current reproduction for the different age groups. Because only N-1 and N-2 were shown to be significant predictors in the model, we tested whether the weight of these factors differed between the age groups. We found that this was the case for N-1 ($\chi^2(3) = 8.58$, $p = 0.035$, $BF_{01} = 17.19$), although the Bayes factor suggests that there was more evidence for the absence of this difference. Post-hoc multiple comparisons showed that the effect of objective N-1 was stronger for 5-year-olds than for adults ($\beta = -0.16$, $z = -3.05$, $p = 0.012$). No other contrasts reached significance (ps > 0.228). There was no difference between age groups for N-2 ($\chi^2(3) = 6.98$, $p = 0.073$, $BF_{01} = 181.36$). In summary, reproductions were significantly influenced by previously presented intervals. In addition, this N-1 effect was stronger for the younger children compared to adults.

3.4.2. Subjective previous durations

Whereas participants might be influenced by recent *objective durations*, it is also possible that they rely on their subjective experience of this objective duration, i.e., their own temporal production (e.g., Schlichting et al., 2018). To test this idea, we again started with the previously established model mentioned in section 3.4.1, and sequentially added previous *subjective durations* (in trial N-1, N-2, N-3, etc.), that is, previous reproductions, to the model. We found that all previous subjective durations up to N-7 contributed significantly to the current reproduction (χ^2s(1) > 18.30, ps < 0.001, $BFs_{01} < 0.01$). We decided that the effect of previous trials beyond N-7 could not be established reliably, because only less than half of the data could be used for these models.

Figure 4B shows the beta weights for the four most recent previous subjective durations for the different age groups. For presentation purposes, we decided to only show the weights up to N-4, nevertheless, a figure showing the weights up to N-7 can be found at https://osf.io/k3znf/.We found that weights of N-3 and N-6 differed significantly between the different age groups ($\chi^2(3) = 11.66$, $p = 0.009$, $BF_{01} = 30.21$ and $\chi^2(3) = 8.94$, $p = 0.030$, $BF_{01} > 100$). However, after adding the random slopes of duration, range, N-1 and N-2, post-hoc multiple comparisons showed that there were no significant differences between the age groups in the effect of N-3 (ps > 0.276). However, the effect of N-6 was larger for 6-year-olds than for 5-year-olds ($\beta = 0.08$, $z = 2.60$, $p = 0.045$). There were no other differences (ps > 0.393).

Although the participants in all age groups might rely on previous subjective durations, this effect could potentially reflect performance drift over the experiment. For example, in certain phases of the experiment, a participant might be less willing to make longer responses compared to other phases. To disentangle the influence of the previous subjective duration from this local performance drift, we calculated the relative error of the reproduction in each trial (error = [reproduced duration - objective duration]/objective duration) (see Schlichting et al., 2018). In the case of performance drift, we would expect that a previous negative error (that is, a too short reproduction) in the previous trial would also lead to negative error in the current trial. In contrast, if the current reproduction depends on the actual previous subjective experience, we would expect that the relative error would reflect the duration of the previous reproduction (that is, a more positive error if the previous reproduction was long and a more negative error if the previous reproduction was long).

Starting with a model with relative error as the dependent variable, the same fixed factors used in section 3.4.1 and subject as a random factor, we alternately added previous reproductions (N-1, N-2, N-3, etc.) and relative error in the previous trials to the model. We found that both the previous reproductions and the previous relative errors up to N-7 improved the model (*p*s < 0.004), indicating that some of the sequential effects can be explained by performance drift, but there was still a significant influence of the actual previous subjective duration.

Figures 4C and 4D show the influence of the relative error and the subjective duration in the four most recent trials on the current reproduction. To test whether the weights differed between the age groups, we sequentially and alternately added the interaction terms of the previous subjective durations and age group, and of previous relative error and age group, to the model. We found that for the effect of previous subjective duration was different for N-1 and N-3 (χ^2s(3) > 8.28, *p*s < 0.041, BFs_{01} < 3.64). In addition, the effect of the previous error in N-1 and N-6 differed between age groups (χ^2s(3) > 8.71, *p*s < 0.044, BFs_{01} < 1.49). Post-hoc multiple comparisons showed that the effect of subjective N-1 was lower for adults than for 5- and 7-year-olds (*p*s < 0.035). For subjective N-3, no contrast reached significance (*p*s > 0.208). Post-hoc comparisons of the effect of relative error in N-1 showed that the effect was lower for 5-year-olds compared to 6-year-olds and adults (*p*s < 0.034). The contrasts also suggested a higher weight for adults compared to 6 and 7-year-olds, but these effects were borderline significant (*p*s < 0.091). No contrast reached significance for the relative error in N-6 (*p*s > 0.192).

To summarize, we found that previous subjective durations influenced the current reproduction, but found no apparent differences between age groups in this respect. However, when we disentangled the influence of previous subjective duration and performance drift, we found adults had a higher influence of performance drift compared to the children. This pattern is reversed when we looked at the weight of previous subjective duration: the children (at least 5 and 7-year-olds) relied more on the previous subjective duration than the adults.

Insert Figure 4 about here

4. Discussion

In our study, children from 5 to 7 years old and adults performed a ready-set-go reproduction task with two different duration distributions. Our results showed an underestimation of reproduced durations as the length of durations increased, especially in young children. This replicated the results found in most studies in children that employ temporal reproduction task (e.g., Droit-Volet et al., 2015; Karaminis et al., 2016; Szelag et al., 2002). This temporal underestimation suggests that factors related to motor impulsivity have likely affected the children's temporal reproductions (Droit-Volet, 2010). This is consistent with the results in rhythmic time interval tasks showing that young children have difficulty in reproducing time intervals far from their Spontaneous Motor Tempo (McAuley et al., 2006; Monier & Droit-Volet, 2016). Children indeed have reduced self-control capacities, and as such, it is difficult for them to inhibit initial response (e.g. the dominant response) (Fox, Henderson, Marshall, Nichols & Ghera, 2005; Klenberg, Korkman & Lahti-Nuuttila, 2001).

This consistent underestimation of long duration might limit the validity of Bayesian modelling, because it is difficult to distinguish between effects coming from the motor component and those resulting from the temporal prior.

Nevertheless, the underestimation bias obtained in our study could be considered in our regression analyses of the age-related differences in the effect of temporal context on performance. The decreased slope of reproductions for children compared to adults provides evidence for a stronger central tendency effect in children. This is in concert with recent studies showing that central tendency effects progressively decrease with age (Sciutti, Burr, Saracco, Sandini & Gori, 2014; Karaminis et al., 2016). Furthemore, we found that the variance in temporal reproduction (as quantified by the coefficient of variation) was higher in all children compared the adults and in the 5-year-olds compared to the 7-year-olds. A higher central tendency effect was thus observed in participants with a lower sensitivity to time. These findings are in line with the idea that the noisier the internal representation of the interval, the larger the central tendency effect will be (Jazayeri & Shadlen, 2010; Acerbi, Wolpert & Vijayakumar, 2012).

In addition, our study suggests that this central tendency effect is due to a greater use of prior presented durations in the experimental session. Indeed, our results showed an effect of global context on temporal reproductions in all age groups: the overlapping duration (0.9 s) was systematically judged longer in the long than in the short context condition. However, despite the noisier reproductions and flatter slopes in the youngest children, we did not find any statistical difference in this global context effect between the age groups. In contrast, our results on the local (trial-by-trial) context effect revealed that the duration presented in the most recent trials had a greater impact on the reproduction of a given duration in the children than in the adults. However, our results revealed that only the most recently presented durations (N-1 and N-2) influenced the participants' time judgments. In sum, the temporal impact of objective duration presented in the previous trial was stronger for 5-year-olds than for adults. If we consider the Bayesian framework, we could thus conclude that, because of a highly noisy percept, the subjective estimation of the younger children is tilted toward previous experiences (the prior) more than it is tilted toward the perceived interval (the likelihood).

As a novel way of looking at the influence of subjective experience, we have not only tested the effect of the objective durations presented on current time judgment, but also that of previous subjective durations, i.e., the participants' own temporal reproduction. We distinguished this effect from general drifts in performance by examining the unique contribution of previous individual reproductions and the previous errors on the current reproduction. We found that both of these factors had a continuing impact (at least up to N-7). However, for the most recent previous trial (i.e., N-1), we found that the effect of both the subjective duration and relative error differed between the age groups. Consistently with the objective duration effect, the children (5 and 7 years) relied more on their previous subjective duration than the adults. Contrariwise, the influence of previous relative error was higher for the adults than for the children, indicating that the reproductions of adults were subject to more reliable performance drifts. These novel findings suggest that, compared to adults, children rely more on the temporal context than on the evaluation of their misjudgement. This is in line with the idea that humans possessearly abilities for statistical learning (Karaminis et al., 2016), since children continuously integrate priors into their current production. These abilities have already been observed in infants and newborns (Kirkham et al., 2007, 2002; Bulf et al., 2011). In contrast, learning from produced errors would emerge in great part later during childhood, explaining the higher performance drift in adults with the development of executive functions, that is, when children become able to evaluate their performance and their evolution during learning. Indeed, among the different aspect of executive functions that develop through childhood, one could notably cite that of error evaluation (Kirkham, Cruess & Diamond, 2003), allowing children to apply knowledge to their own behaviour.

In summary, our results demonstrated that the central tendency effect in temporal reproduction is stronger in children than in adults, and that children's current temporal reproductions rely more on durations presented in recent trials. This finding can be linked to the children's noisier representation of time. Consistent with Bayesian theory, a noisy timing system led participants to further base their estimation on the previous experiences rather than on the perceived stimulus. However, the influence of relative error (subjective produced duration) was higher for the adults than for the children. This new finding suggests that, unlike adults, children rely to a greater extent on the temporal context than on the evaluation of their misjudgement. Future studies might further investigate whether the influence of context in temporal judgment in children generalizes to different contexts and temporal tasks.

References

Acerbi, L., Wolpert, D. M., & Vijayakumar, S. (2012). Internal representations of temporal statistics and feedback calibrate motor-sensory interval timing. *PLoS computational biology*, *8*(11), e1002771.

Adams, W. J., & Mamassian, P. (2004). Bayesian combination of ambiguous shape cues. *Journal of vision*, *4*(10), 7-7.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R package version*, *1*(7).

Battaglia, P. W., Jacobs, R. A., and Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *J. Opt. Soc. Am. A. Opt. Image Sci. Vis.* 20, 1391–1397.

Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, *121*(1), 127-132.

Brainard, D. H. (1997). The Psychophysics Toolbox, *Spatial Vision 10*:433-436.

Cicchini, G. M., Arrighi, R., Cecchetti, L., Giusti, M., & Burr, D. C. (2012). Optimal encoding of interval timing in expert percussionists. *Journal of Neuroscience*, *32*(3), 1056-1060.

Crowder, A. M., & Hohle, R. H. (1970). Time estimation by young children with and without informational feedback. *Journal of Experimental Child Psychology*, *10*(3), 295-307.

Coull, J., & Droit-Volet (In press). Explicit understanding of duration develops implicitly through action. *Trends In Cognitive Sciences.*

Damsma, A., van der Mijn, R., & van Rijn, H. (2018). Neural markers of memory consolidation do not predict temporal estimates of encoded items. *Neuropsychologica, 117*, 36-45. doi:10.1016/j.neuropsychologia.2018.04.039

Di Luca, M., & Rhodes, D. (2016). Optimal perceived timing: Integrating sensory information with dynamically updated expectations. *Scientific reports*, *6*, 28563.

Droit-Volet, S. (2010). Stop using time reproduction tasks in a comparative perspective without further analyses of the role of the motor response: The example of children. *European Journal of Cognitive Psychology*, *22*(1), 130-148.

Droit-Volet, S. (2013). Time perception in children: A neurodevelopmental approach. *Neuropsychologia*, *51*(2), 220-234.

Droit-Volet, S. (2016). Development of time. *Current Opinion in Behavioral Sciences*, *8*, 102-109.

Droit-Volet, S., Wearden, J. H., & Zélanti, P. S. (2015). Cognitive abilities required in time judgment depending on the temporal tasks used: a comparison of children and adults. *The Quarterly Journal of Experimental Psychology*, *68*(11), 2216-2242.

Dyjas, O., Bausenhart, K. M., & Ulrich, R. (2012). Trial-by-trial updating of an internal reference in discrimination tasks: Evidence from effects of stimulus order and trial sequence. *Attention, Perception, & Psychophysics*, *74*(8), 1819-1841.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429.

Fox, N. A., Henderson, H. A., Marshall, P. J., Nichols, K. E., & Ghera, M. M. (2005). Behavioral inhibition: linking biology and behavior within a developmental framework. *Annu. Rev. Psychol.*, *56*, 235-262.

Hallez, Q., & Droit-Volet, S. (2017). High levels of time contraction in young children in dual tasks are related to their limited attention capacities. *Journal of experimental child psychology*, *161*, 148-160.

Hollingworth, H. The central tendency of judgment. J. Phil. Psychol. Sci. Meth. 7, 461–469 (1910).

Hothorn, T., Bretz, F., Westfall, P., Heiberger, R.M. & Schuetzenmeister, A. (2013) Multcomp: Simultaneous Inference in General Parametric Models R Package Version 1.2 ‐ 18. http://cran.r-project.org/web/packages/multcomp/index.html.

Indraccolo, A., Spence, C., Vatakis, A., & Harrar, V. (2016). Combined effects of motor response, sensory modality, and stimulus intensity on temporal reproduction. *Experimental brain research*, *234*(5), 1189-1198.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature neuroscience*, *13*(8), 1020.

Karaminis, T., Cicchini, G. M., Neil, L., Cappagli, G., Aagten-Murphy, D., Burr, D., & Pellicano, E. (2016). Central tendency effects in time interval reproduction in autism. *Scientific reports*, *6*.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35-B42.

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.

Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension ‐ switching task. *Developmental Science*, *6*(5), 449-467.

Kleiner M, Brainard D, Pelli D, 2007, “What’s new in Psychtoolbox-3?” Perception 36 ECVP Abstract Supplement.

Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2001). Differential development of attention and executive functions in 3-to 12-year-old Finnish children. *Developmental neuropsychology*, *20*(1), 407-428.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS one*, *2*(9), e943.Lapid, E., Ulrich, R., & Rammsayer, T. (2008). On estimating the difference limen in duration discrimination tasks: A comparison of the 2AFC and the reminder task. *Perception & psychophysics*, *70*(2), 291-305.

Lejeune, H., & Wearden, J. H. (2009). Vierordt's The Experimental Study of the Time Sense (1868) and its legacy. *European Journal of Cognitive Psychology*, *21*(6), 941-960.

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *69*(1), 1-33.

Mamassian, P., Landy, M., & Maloney, L. T. (2002). Bayesian modelling of visual perception. *Probabilistic models of the brain*, 13-36.

McAuley, J. D., & Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: a comparison of interval and entrainment approaches to short-interval timing. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(6), 1102.

McAuley, J., Jones, M., Holub, S., Johnston, H., & Miller, N. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*(3), 348–367.

Miyazaki, M., Nozaki, D., & Nakajima, Y. (2005). Testing Bayesian models of human coincidence timing. *Journal of neurophysiology*, *94*(1), 395-399;

Monier, F., & Droit-Volet, S. (2018, first publication 2016). Synchrony and Emotion in Children and Adults. *International Journal of Psychology*, *53*, 3, 184-193.

Morey, R. D., Rouder, J. N., & Jamil, T. (2014). BayesFactor: Computation of Bayes factors for common designs. *R Package Version*, *0.9*(8).

Petzschner, F. H., Maier, P., & Glasauer, S. (2012). Combining symbolic cues with sensory input and prior experience in an iterative Bayesian framework. *Frontiers in integrative neuroscience*, *6*.doi

Rattat, A. C., & Tartas, V. (2017). Temporal categorization of familiar actions by children and adults. *Timing & Time Perception*, *5*(1), 61-76.

Rhodes, D. (2018). On the Distinction Between Perceived Duration and Event Timing: Towards a Unified Model of Time Perception. *Timing & Time Perception*, *6*(1), 90-123.

Schlichting, N., Damsma, A., Aksoy, E. E., Wächter, M., Asfour, T., & van Rijn, H. (2018). Temporal Context Influences the Perceived Duration of Everyday Actions: Assessing the Ecological Validity of Lab-Based Timing Phenomena. *Journal of Cognition*, *1*(1).

Sciutti, A., Burr, D., Saracco, A., Sandini, G. & Gori, M. Development of context dependency in human space perception. Exp. Brain Res. 232, 3965–3976 (2014).

Shi, Z., & Burr, D. (2016). Predictive coding of multisensory timing. *Current opinion in behavioral sciences*, *8*, 200-206.

Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature neuroscience*, *9*(4), 578.

Szelag, E., Kowalska, J., Rymarczyk, K., & Pöppel, E. (2002). Duration processing in children as determined by time reproduction: implications for a few seconds temporal window. *Acta Psychologica*, *110*(1), 1-19.

Taatgen, N., & van Rijn, H. (2011). Traces of times past: representations of temporal intervals in memory. *Memory & cognition*, *39*(8), 1546-1560.

van Rijn, H. (2016). Accounting for memory mechanisms in interval timing: A review. *Current Opinion in Behavioral Sciences* 8: 245–249.

Verstynen, T., & Sabes, P. N. (2011). How each movement changes the next: an experimental and theoretical study of fast adaptive priors in reaching. *Journal of Neuroscience*, *31*(27), 10050-10059.

Vierordt, K. Der Zeitsinn nach Versuchen (Laupp, Tubingen, Germany, 1868).

Walsh, V. (2003) A theory of magnitude: common cortical metrics of time, space and quantity. Trends Cogn. Sci. 7, 483–488 13

Wearden, J. (2016). *The psychology of time perception*. Springer.

Funding: This study was funded by H2020 European Research Council (BE) (TIMESTORM - H2020- FETPROACT-2014).

Acknowlegement: The authors thank the directors and the professors (D. Avond, S. Perrier, C. Garcia, J. Joyaux, C. Aymard, L. Guimoneau, E. Parillaud, E. Bernot, S. Sanchez-Silva, A. Le Mouroux, S. SanchezSilva, S. Dompnier) from the Primary Schools Philippe Arbos and

Victor Hugo as well as the Elementary Schools Denis Diderot and Nestor Perret, all located in Clermont-Ferrand. Thanks also to Warrick Roseboom for his help on this project.

Figure captions:

Figure. 1. Ready-set-go procedure: (a) temporal context, (b) procedure.

Figure. 2. Average reproductions of the durations (A) and CV value for the different age groups. Error bars represent the standard error of the mean.

Figure. 3. Average difference of the 0.9 s reproduction between the short and the long context for the different age groups. Error bars represent the standard error.

Figure. 4. The weight of previous durations as quantified by the beta estimates of our linear mixed models. Figure A shows the effect of previous *objective* duration on the current reproduction, whereas figure B shows the effect of previous *subjective* duration on the current reproduction. To disentangle performance drift from the effect of previous subjective duration, Figure C and D shows the weights of the previous relative error and previous reproduction on the current relative error, respectively.