Cognition 257 (2025) 106078

Contents lists available at ScienceDirect

Cognition

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

Uncovering the latent structure of human time perception

Renata Sadibolova [a,b,c,*], Curtis Widmer [a], Zoe Fletcher [a], Soraya Weill [d], Devin B. Terhune [a,b,**]

[a] Department of Psychology, Goldsmiths, University of London, London, UK
[b] Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[c] School of Psychology, University of Roehampton, London, UK
[d] Department of Experimental Psychology, University of Oxford, Oxford, UK

ARTICLE INFO

Keywords:
Central tendency effect
Temporal reproduction
Subsecond and suprasecond timing
Changepoint analysis
Bifactor analysis

ABSTRACT

One of the ongoing controversies in interval timing concerns whether human time perception relies on multiple distinct mechanisms. This debate centres around whether subsecond and suprasecond timing may be attributed to a single semi-uniform timing system or separate and interacting cognitive systems. Whereas past studies offer valuable insights, this study overcomes previous limitations by adopting multiple convergent statistical approaches in a design with strong statistical power. We conducted two online experiments involving participants reproducing temporal intervals ranging from 400 to 2400 ms (Experiment 1; $N = 302$) and 1000 to 2000 ms (Experiment 2; $N = 302$). We contrasted the application of exploratory factor analysis and structural equation modelling to differentiate distinct latent structures underlying duration reproduction patterns. Additionally, we compared the model outcomes with results from changepoint analysis models fitted to individual participants' data. In both experiments, these analyses yielded evidence for a two-factor model comprising a general timing factor spanning the full interval range and a second factor capturing the regression to the mean of presented stimulus intervals (central tendency bias). We observed a low proportion of detected changepoints, further supporting the limited evidence for a hypothesized discontinuity between distinct underlying systems, while also finding that changepoint detection patterns were predicted by latent factor scores. These results suggest that the central tendency bias should be considered when investigating potential discontinuities in interval timing systems. Our work contributes to the integration of factor analytic and computational modelling approaches in the study of time perception and has implications for the measurement and interpretation of interval timing in a range of contexts.

1. Introduction

Temporal information is processed across multiple timescales, each uniquely implicated in distinct perceptual and behavioural functions. For instance, motor actions and speech generation are typically executed within brief subsecond intervals whereas decision making tends to span suprasecond intervals (Brody, 2003; Edwards, Alder, and Rose, 2002; Merchant and Georgopoulos, 2006; Schirmer, 2004; Sohn and Carlson, 2003). These observations have contributed to debates about whether the processing mechanisms for subsecond and suprasecond time perception are similar or different, and regarding the nature of their continuity for the latter case (Nani et al., 2019; Rammsayer and Pichelmann, 2018). This raises important questions about whether these processes are consistent across different durations or whether distinct timing systems govern the perception of short versus long intervals. This study adopts a data-driven approach, seeking to contribute to these debates by investigating the latent structure of temporal performance.

Multiple studies suggest that semi-distinct timing systems exist for intervals in the millisecond and second ranges, with different systems being activated as stimulus duration increases (Buhusi and Meck, 2005; Gibbon, Malapani, Dale, and Gallistel, 1997; Gooch, Wiener, Hamilton, and Coslett, 2011). In support of these propositions, pharmacological disruption of working memory selectively impairs suprasecond interval discrimination, whereas different pharmacological agents selectively alter timing in the subsecond range (Rammsayer, 1993, 2008). Similar double dissociations have been associated with different dopamine

* Corresponding author at: School of Psychology, Roehampton Ln, London SW15 5PH, UK.
** Correspondence to: Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, 16 De Crespigny Park, London SE5 8AB, UK.
E-mail addresses: renata.sadibolova@roehampton.ac.uk (R. Sadibolova), devin.terhune@kcl.ac.uk (D.B. Terhune).

https://doi.org/10.1016/j.cognition.2025.106078
Received 8 July 2023; Received in revised form 19 December 2024; Accepted 30 January 2025
Available online 11 February 2025

genes (Wiener, Lohoff, and Coslett, 2011) and behavioural tasks that target either short or long intervals (Rammsayer and Lima, 1991; Rammsayer and Ulrich, 2001). Consistent with this evidence, functional neuroimaging findings have further implicated different brain regions for brief intervals in the milliseconds range and for longer, cognitively-mediated intervals (Mondok and Wiener, 2023; Rammsayer and Pichelmann, 2018).

Citing this research, Rammsayer and Pichelmann (2018) proposed a model with modality-specific neurocognitive mechanisms subserving the timing of "sensory-automatic" intervals (below ~100–500 ms; Buonomano, Bramen, and Khodadadifar, 2009; Sadibolova, Sun, and Terhune, 2021) and longer intervals supported by a more cognitive, higher-order system (see also Lewis and Miall, 2003; Lusk, Petter, and Meck, 2020; Mondok and Wiener, 2023). The model assumes a transitional stage with system overlap, where the influence of the sensory-automatic system gradually diminishes as intervals lengthen. A purported demarcation point between these timing systems in the literature has been 1000 ms (Bangert, Reuter-Lorenz, and Seidler, 2011; Buhusi and Meck, 2005; Gooch, Wiener, Hamilton, and Coslett, 2011) although some research has suggested it to be ~500 ms (Rammsayer and Lima, 1991), ~1200 ms (Grondin, Meilleur-Wells, and Lachance, 1999; Grondin, Ouellet, and Roussel, 2004), ~1500 ms (Gibbon, Malapani, Dale, and Gallistel, 1997; Grondin, 2012), or no earlier than ~3–4 s (Lewis and Miall, 2009). To clarify this ambiguity, our study examines an interval range that exceeds the boundary proposed by Rammsayer and Pichelmann (2018) but remains within the upper limit suggested by Lewis and Miall (2009). Based on mixed evidence for this interval range, we may expect to find evidence for a uniform timing mechanism or distinct overlapping timing systems.

Temporal performance can provide valuable insights into underlying timing systems. Researchers have traditionally assessed either subjective time (e.g., stimulus interval reproduction times) or the coefficient of variation (CV), a measure of standardised variability across reproduced intervals (Buhusi and Meck, 2005; Grondin, 2012; Lewis and Miall, 2009). These approaches are based on principles of scalar expectancy theory (SET; Gibbon, 1977) according to which subjective time estimates increase steadily with objective intervals, and the variability in these estimates scales with interval length and follows a constant trend. The rationale behind these studies is that sudden changes in these measures across the investigated interval range may indicate a shift to a different timing system. Our experiments focus on reproduced intervals and omitted CV analyses because previous research found that the CV was not constant, showing logarithmically decreasing trends even within small interval ranges that were expected to rely on a single timing system (Lewis and Miall, 2009).

Several studies have applied statistical techniques such as principal component analysis, factor analysis, and structural equation modelling to investigate the features of and an overlap between the putative timing systems (Rammsayer and Altenmüller, 2006; Rammsayer and Brandler, 2004; Rammsayer and Troche, 2014; Stauffer, Haldemann, Troche, and Rammsayer, 2012). Other approaches, such as the use of hierarchical clustering methods and computation of cross-task correlations to investigate distinctions in timing systems across tasks (Bangert, Reuter-Lorenz, and Seidler, 2011; Merchant, Zarco, Bartolo, and Prado, 2008), have complemented efforts to explore the latent structure of timing mechanisms. Some found evidence suggesting that there might be two related systems at work, while others pointed to a single, unified system. However, the factor-analytical studies were often limited by small sample sizes, which may have hindered their ability to accurately capture the latent structure of interval timing. In this study, we aimed to apply multiple complementary latent variable modelling approaches to characterize the latent structure of interval timing.

In two well-powered experiments (*N*s > 300), participants completed visual reproduction tasks spanning 400-2400 ms (Experiment 1) and 1000-2000 ms (Experiment 2). We applied exploratory factor analysis (EFA) and exploratory structural equation modelling (ESEM) to reproduction times to contrast competing models of the latent structure of temporal performance in order to determine whether timing performance was subserved by a single, or multiple, timing systems and whether their recruitment scaled with stimulus intervals. We further supplemented these analyses with subject-level changepoint detection, commonly used in time-series signal analysis to detect shifts in system behaviour. We expected these complementary techniques to provide converging evidence for these competing predictions and, in the case of distinct timing systems, to explicate the transition between them.

2. Methods

2.1. Participants

Although no formal a priori sample size estimation was performed, we planned to include a sample size of $N = 300$, which allows for a sample-to-variable ratio of >10, which is considered as adequate for factor analysis (Osborne and Costello, 2004). In *Experiment 1*, 473 individuals initiated the experiment, but 145 did not complete it. Of the remaining 328 participants, 26 were removed as outliers (see Supplemental Materials), leaving 302 for the final analyses (age: $M = 33.10$, $SE = 0.70$; gender: female: $n = 175$, male: $n = 126$, other: $n = 1$). In *Experiment 2*, 402 individuals started the experiment, with 344 participants completing it. After removing 42 outliers (see Supplemental Materials), the final sample size was 302 (age: $M = 36.85$, $SE = 0.72$; gender: female: $n = 169$; male: $n = 132$; other: $n = 1$). Participants were recruited via Prolific (https://www.prolific.co/) with the inclusion criteria being 18–65 years-old, native English speaker and a Prolific approval rate of ≥95 %. All participants provided informed consent in accordance with local ethical approval and were compensated at a £5 hourly rate. The data used in this study is available at osf.io/gm9wv (Sadibolova & Terhune, 2024).

2.2. Materials

The task was executed using custom-written JavaScript code accessed via a URL on the Prolific platform. Whereas the experiment's online nature meant that participants used different viewing setups, smartphones and tablets were excluded and a standard keyboard was required for responses. No restrictions were placed on browser choice, machine specifications, or internet speed. Stimulus durations were aligned with a typical monitor refresh rate of 60 Hz.

2.2.1. Temporal reproduction task

A temporal reproduction task was modelled after previously used paradigms (Grondin, 2010). Trials began with a variable interstimulus interval (ISI) of 675, 725, 775, or 825 ms, followed by the presentation of a blue circle stimulus (ø ~3 cm). In *Experiment 1*, 11 stimulus intervals were used, ranging from 400 to 2400 ms in 200 ms increments. In *Experiment 2*, 14 intervals were logarithmically spaced from 983 to 2000 ms. After a second ISI (fixed 500 ms), participants were prompted with a fixation cross to reproduce the stimulus interval by holding down the spacebar.

2.2.2. Procedure

After reading the information sheet and providing informed consented, participants provided their demographic information and received instructions for completing the task. They were instructed to focus on the centre of the monitor at all times and to avoid using strategies such as imagination, repetitive movements, humming, or counting, as these could influence timing performance (Grondin, Ouellet, and Roussel, 2004). In *Experiment 1*, participants completed one block of 11 practice trials followed by 4 blocks of 44 experimental trials (four of each interval), totalling 176 trials. In *Experiment 2*, participants completed a block of 14 practice trials and 4 blocks of 42 experimental trials (three of each interval), totalling 168 trials. Neither study included

feedback for either the practise or experimental trials.

2.3. Statistical analyses

The analyses were performed in MATLAB 2019b (MathWorks, Natick, MA), MPLUS Version 7.3 (Muthén and Muthén, 2007) and R software (R Core Team, 2021).

2.3.1. Data processing

We computed the median reproduced intervals for each stimulus and participant, as medians are less susceptible than means to noise and outliers that are significantly more prevalent in online data collection (Chmielewski and Kucker, 2020; Peer, Rothschild, Gordon, Evernden, and Damer, 2021). In *Experiment 1*, we examined the data for deviant linear trends across reproduced intervals and excluded 26 participants due to poor performance or suspected equipment malfunction. We then removed for each participant and stimulus any outliers, i.e., reproduced durations that surpassed a three-median absolute deviation cut-off (*M*: 3.3 %, *SE*: 0.1, range: 0–18.8 %). In *Experiment 2*, we adopted a similar strategy but adjusted for logarithmically spaced stimuli. A second-order polynomial model was fitted to the median reproduced intervals, and outliers were identified using the coefficients from this model. Forty-two participants identified as outliers with the MATLAB's *robustcov* algorithm were excluded. Using a three-median absolute deviation cut-off, we also removed outlier trials for each participant and stimulus (*M*: 3.2 %, *SE*: 0.1, range: 0–11.1 %). Further details are provided in Supplemental Materials Section 1.

2.3.2. Dimensionality assessment

To assess the underlying structure of temporal reproduction performance, we first applied EFA to median reproduced intervals in each experiment. The ESEM was planned and implemented only when the EFA yielded a correlated-factors model (Fig. 1) as the best fitting solution for the data. This was necessary because the correlation can obscure the unique variance each factor represents (Kline, 2023), complicating the identification of constructs involved (e.g., distinct timing systems). We planned to use ESEM to assess two models addressing this issue. First, a second-order factor model, which seeks to explain the correlation between the factors through a higher-order construct. Second, a bifactor model, which isolates the shared variance in a general factor, allowing specific factors to be defined by their unique variance (Kline, 2023; Markon, 2019).

We used standard factor-extraction approaches in the EFA to identify the latent structure of reproduced time intervals. These included parallel analysis (Horn, 1965), which retains only factors with eigenvalues exceeding those derived from randomly generated data. We also used exploratory graph analysis (EGA; Golino and Epskamp, 2017) due to its advantages in models with highly correlated factors. The Supplemental Materials Section 2 provides additional details, including the minimum average partial (MAP; Velicer, 1976) and very simple structure (VSS; Revelle and Rocklin, 1979) statistics for models with varying factor numbers. Since the correlated-factors model provided the best fit in both experiments, we proceeded with ESEM.

Our next aim was to determine whether more advanced ESEM models could account for the covariance among the first-order factors. However, second-order factor modelling was not possible for the data in either experiment, as it requires at least three first-order factors to synthesize a broader overarching construct (Kline, 2023; Muthén and Muthén, 2007). With only two correlated factors in our final EFA model (Tables 1 and 2), the second-order factor model could not be identified and would have merely restated the factor correlation without offering new insights. By contrast, bifactor modelling (Kline, 2023; Markon, 2019) allows for the orthogonalization of the correlated factors by attributing shared variance to a general factor and identifying each specific factor based on its unique variance. We assessed multiple exploratory bifactor models, each with a different number of specific orthogonal factors alongside the general factor. The model with the best fit statistics was selected as the ESEM outcome.

Model fit was compared using conventional criteria (Weston and Gore, 2006): root mean square error of approximation (RMSEA) < 0.10, comparative fit index (CFI) and Tucker-Lewis index (TLI) > 0.90, and the standard root mean square residual (SRMR) < 0.10 for adequate model fit. Exceptionally good model fit cut-off values are <0.06, > 0.95 and < 0.08, respectively. Additional bifactor model indices (Rodriguez, Reise, and Haviland, 2016) indicate the strength of unidimensionality vs. multidimensionality (Explained Common Variance; ECV and Percentage of Uncontaminated variance; PUC), strength of individual factors (omega indices; ω), whether the factor scores reflect replicable individual differences (Factor Determinacy index; FD), and the construct replicability (H). A general factor with PUC > 0.80 and ECV > 0.70 implies a strong unidimensional model with a relative bias <10 % (Bonifay, Reise, Scheines, and Meijer, 2015) and individual variables

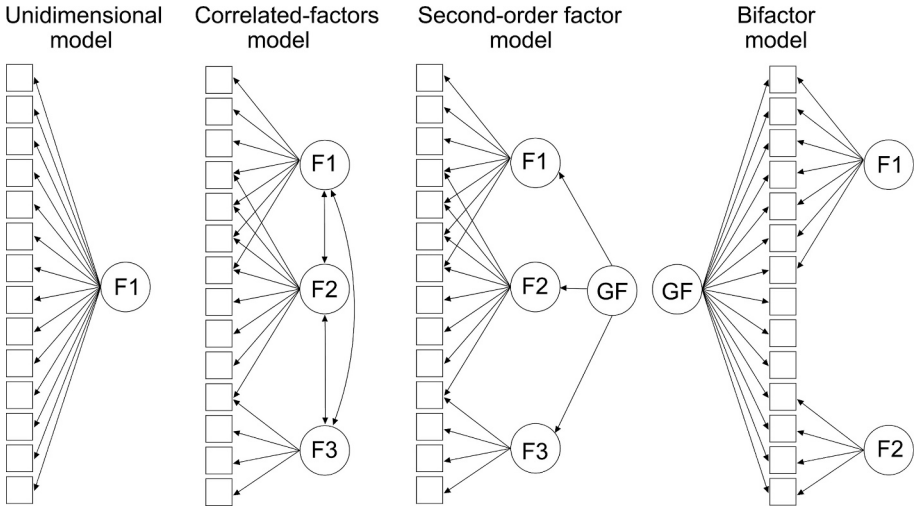


Fig. 1. Schematic depiction of tested latent variable models applied to reproduced durations in a visual temporal reproduction task. Latent variables are represented by circles and the observed variables by squares. Single-headed arrows represent a causal path whereas double-headed arrows denote correlation. The *unidimensional model* is characterised by all items loading on a single factor. The *correlated-factors model* allows for two or more correlated factors, each with strong loadings for a subset of items. The *Second-order-factor model* includes an additional higher-order general factor that accounts for covariation among the first-order factors. The *bifactor* model attributes shared item variance to a first-level GF and the remaining orthogonal variance to specific first-order factors. F = factor; GF = general factor.

Table 1
Fit statistics for models applied to reproduced intervals in Experiment 1 ($N = 302$), including the indices specific to bifactor modelling.

Model	Model fit indices				Bifactor model indices				
	RMSEA [CI]	CFI	TLI	SRMR	H	FD	ω	ECV	PUC
Unidimensional	0.18 [0.17, 0.19]	0.72	0.65	0.15					
Correlated-factors	0.09 [0.07, 0.10]	0.95	0.92	0.02					
Bifactor	**0.09 [0.07, 0.10]**	**0.95**	**0.92**	**0.02**					**0.00**
G					**0.72**	**0.94**	**0.01**	**0.21**	
F1					**0.96**	**0.98**	**0.99**	**0.79**	
Bifactor	0.18 [0.16, 0.20]	0.85	0.66	0.02					−1.00
G					0.73	0.95	0.02	0.21	
F1					0.96	0.99	0.97		
F2					0.24	0.77	0.01		

Notes. The bolded model demonstrates the best fit while being most parsimonious. G = general factor; F1 = first-order (specific) factor; F2 = second (specific) factor; CI = 95 % confidence intervals; H = Construct replicability index; FD = Factor determinacy index; ω = Omega; ECV = Explained common variance; PUC = Percentage of uncontaminated variance; RMSEA = Root mean square error of approximation; CFI = Comparative fit index; TLI = Tucker-Lewis index; SRMR = Standard root mean square residual.

Table 2
Fit statistics for models applied to reproduced intervals in Experiment 2 ($\boldsymbol{N} = 302$), including the indices specific to bifactor modelling.

Model	Model fit indices				Bifactor model indices				
	RMSEA [CI]	CFI	TLI	SRMR	H	FD	ω	ECV	PUC
Unidimensional	0.15 [0.14, 0.16]	0.89	0.87	0.06					
Correlated-factors	0.03 [0.00, 0.05]	0.99	0.99	0.01					
Bifactor	**0.03 [0.00, 0.05]**	**0.99**	**0.99**	**0.01**					**0.00**
G					**0.98**	**0.99**	**1**	**0.92**	
F1					**0.5**	**0.91**	**0**	**0.08**	
Bifactor	0.04 [0.01, 0.06]	0.99	0.99	0.01					−1.00
G					0.98	0.99	0.99	0.91	
F1					0.51	0.91	0	0.08	
F2					0.11	0.64	0	0.01	

Notes. The bolded model demonstrates the best fit while being most parsimonious. G = general factor; F1 = first-order (specific) factor; F2 = second (specific) factor; CI = 95 % confidence intervals; H = Construct replicability index; FD = Factor determinacy index; ω = Omega; ECV = Explained common variance; PUC = Percentage of uncontaminated variance; RMSEA = Root mean square error of approximation; CFI = Comparative fit index; TLI = Tucker-Lewis index; SRMR = Standard root mean square residual.

with ECV (IECV) > 0.85 are essentially unidimensional (Stucky and Edelen, 2015). A general factor's relative ω reflects reliable variance independent of specific factors, while the ω for specific factors indicates their variance after accounting for the general factor. Recommended values for H and FD are >0.80 and > 0.90, respectively (Gorsuch, 2013; Hancock and Mueller, 2001).

All models used maximum likelihood estimation with a robust option (MLR) in MPLUS (Muthén and Muthén, 2007). This effectively handles violations of multivariate normality (Wang and Wang, 2012), which were observed in the data (Mardia's tests $p < .05$; Mardia, 1970). The models were approximated using the GEOMIN factor rotation method with a bifactor model using the Jennrich-Bentler Bi-GEOMIN orthogonalization (Jennrich and Bentler, 2011). All models were fitted with Epsilon parameter 0.01 and 1000 random starts (Hattori, Zhang, and Preacher, 2017).

2.3.3. Changepoint analysis

Since the EFA and ESEM models estimate uniform parameter values across participants, we subsequently assessed the evidence for changepoints in each participants' median reproduction times across the stimulus intervals in each experiment. In a time-series signal, a change in system behaviour or a shift from one system to another is typically detected as an instance where statistical properties of the signal begin to differ. This may be expressed as a change in direction (trend-based), variance (scale-based), or the cumulative mean (mean-based), as shown in Supplemental Fig. S2.

We applied these different complementary changepoint analyses to reproduced intervals using MATLAB's *findchangepts* function, and *changerob* (Dehling, Fried, and Wendler, 2020) and *bcp* (Erdman and Emerson, 2007) *R* functions. The *findchangepts* algorithm searches for deviations in linear trends and identifies changepoints based on the improvement in residual error of the model. Allowing for variance- and mean-based changepoint detection ('scale' and 'location' options, respectively), the *changerob* algorithm offers an assumption-free approach with changepoints tested for statistical significance at $\alpha = 0.05$. Finally, we complemented these frequentist methods with a Bayesian mean-based changepoint analysis (Barry and Hartigan, 1993), implemented via the *bcp* function, which allowed us to estimate the probability of a changepoint at each stimulus interval. The changepoint algorithms varied in their input requirements. Some were limited to vector inputs with median intervals, while others allowed analysis of the full distribution to detect scale-based patterns.

To assess whether the observed number of participants with detected changepoints was significant, we applied a binomial test. Our aim was to determine, for each changepoint detection approach, whether the number of participants with changepoints was significantly higher than expected, assuming a 5 % probability of detection errors. Binomial test *p*-values smaller than 0.05 suggest that the observed number of changepoints is significant. The results were significant for all but the trend-based analysis in Experiment 1. This suggests that the detection methods are identifying changepoints at rates beyond random detection errors.

Additional analyses included three Fisher's exact tests of association to examine whether the odds of identifying changepoints with one method were associated with those of another. This allowed us to assess the comparability of results across different detection algorithms. These tests were applied to the three frequentist methods only, as the Bayesian approach provides a probability estimate for each interval rather than a detected changepoint(s).

2.3.4. Synthesis of analytical approaches

One of the key objectives of these experiments was to determine whether the two distinct analytical approaches align in identifying the underlying timing systems underlying patterns in temporal reproduction performance. To investigate this, we performed logistic regression analyses using the factor scores obtained from the final ESEM as predictors, and dichotomous changepoints (0 = absent, 1 = present) from each frequentist changepoint analysis as dependent measures. Significant positive coefficients would suggest that individuals with higher factor scores (indicating a closer alignment with the latent constructs identified by ESEM) are more likely to exhibit changepoints in their temporal reproduction performance, whereas significant negative coefficients would imply that those with higher factor scores are less likely to show changepoints. Such results would highlight the robustness of our findings, demonstrating that these analytical methods can jointly reveal consistent insights about the timing systems for temporal reproduction.

3. Results

3.1. Experiment 1

Temporal reproduction patterns showed increasing trends across objective stimulus intervals and larger inter-participant variability for longer intervals (Fig. 2), consistent with findings from previous studies (Grondin, 2010; Zakay, 1990). A noticeable change in the trend around the 1000 ms mark could suggest a shift to a distinct timing system. However, there's another possibility. A common phenomenon in timing performance is the *central tendency bias*, where the reproduction of both short and long intervals typically shifts toward the mean of the stimulus interval range (Acerbi, Wolpert, and Vijayakumar, 2012; Lejeune and Wearden, 2009; Sadibolova and Terhune, 2022). The plateauing trend in Fig. 2 might reflect this bias, weakened by a reduced over-reproduction of short intervals. Central tendency bias can be explained by Bayesian inference, where priors (mental predictions based on past experiences) regularize performance, particularly when the task is more difficult (Sadibolova and Terhune, 2022). Thus, the alternative interpretation of the trend in Fig. 2 is that temporal priors may be exerting a stronger influence on longer interval responses due to higher uncertainty, as predicted by Weber's law. Our next objective was to determine which of these interpretations aligns more closely with the latent structure of these data.

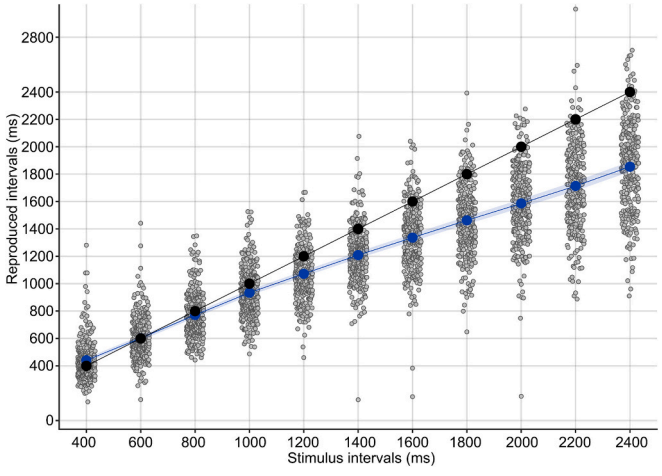


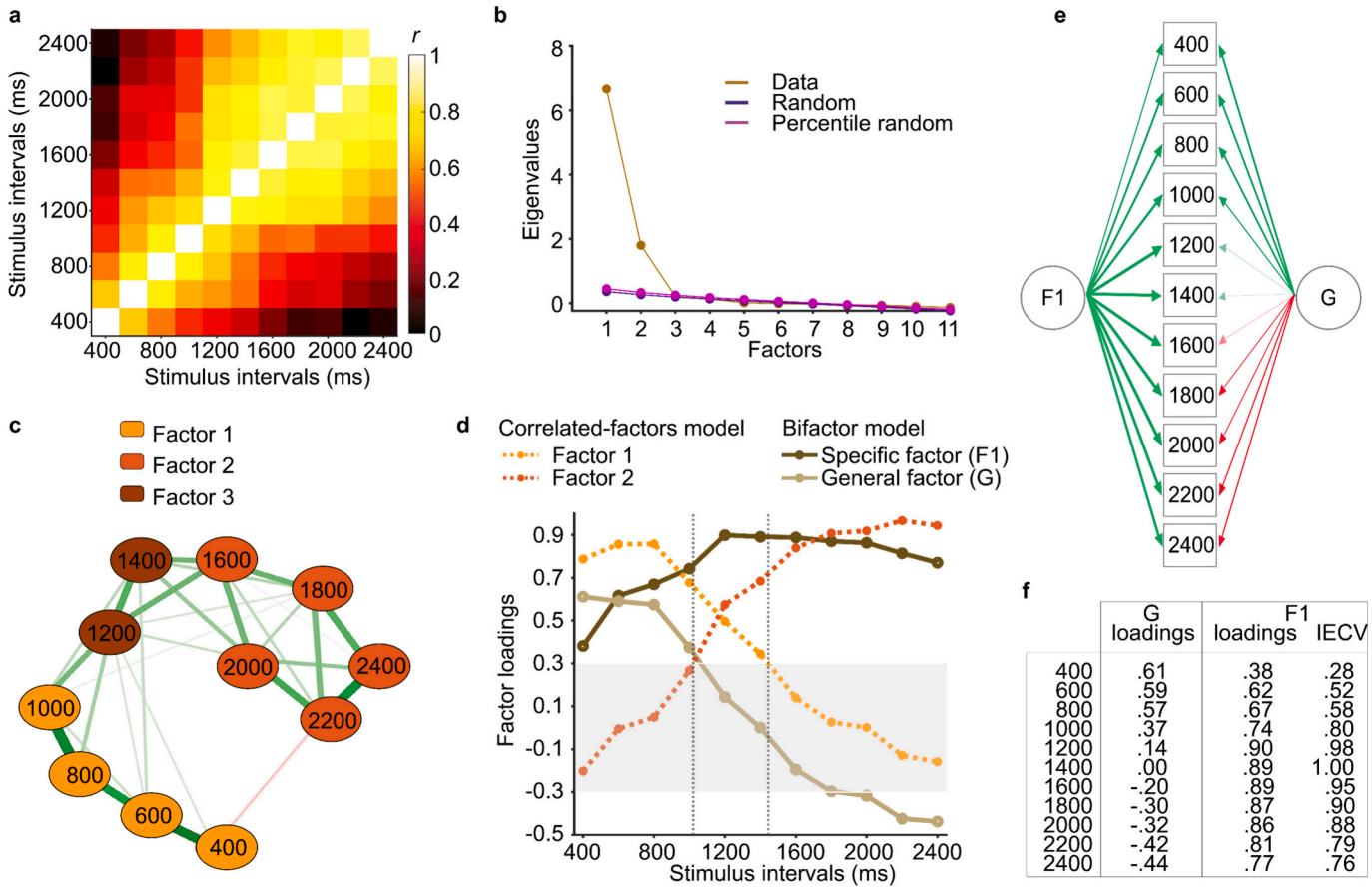
Fig. 2. Reproduced intervals in a visual temporal reproduction task in Experiment 1 ($N = 302$). Grey circles represent median reproduced intervals for each participant for each stimulus interval. Blue circles and the blue shaded area represent sample medians and 95 % confidence intervals, respectively. Black circles denote veridical stimulus intervals and are included for reference. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.1.1. Dimensionality assessment

The Kaiser measure of sampling adequacy ($KMO = 0.92$) and Bartlett test, χ^2 (55) = 3532.40, $p < .001$, indicated that the factor analysis was appropriate. The correlation matrix (Fig. 3a) showed clustering of reproduced intervals for short and long stimuli due to higher correlations for close-range intervals, with those for the longer stimuli being notably larger. This pattern is highly suggestive of two underlying timing systems, which was initially supported by extracting two factors using parallel analysis (Fig. 3b) and based on the MAP and VSS2 statistics (see Supplemental materials Section 2). Although a three-factor solution was identified by the exploratory graph analysis (Fig. 3c), it was not retained given the convergence on the two-factor solution with the remaining factor extraction techniques.

The model fit indices for the correlated-factors solution suggested an adequate fit, superior to the fit of the unidimensional model (Table 1). The factor structure showed short intervals loading on one factor and long intervals on another, with convergence in the mid-stimulus range (1000–1400 ms) as indicated by their cross-loadings (correlated-factors model; Fig. 3d). The factors were moderately correlated, $r(300) = 0.44$, $SE = 0.07$, $p < .001$, which makes interpretation challenging. In our experiment, this outcome may indicate two distinct but overlapping timing systems for short and long intervals, or it may reflect the central tendency effect, capturing the overestimation of short intervals and underestimation of long ones, rather than two separate systems. One approach to address this issue is to use ESEM, which seeks to account for the covariance among the first-order factors.

Although included in our analysis plan, the second-order factor model could not be identified, as it requires at least three first-order factors to capture their correlation within a higher-order overarching construct (see Methods). In contrast, bifactor modelling, which separates the shared variance into a general factor and distinguishes specific factors by their unique variance, can resolve the ambiguity in the correlated-factors model by determining whether the orthogonalized factors reflect a central tendency effect or distinct systems for short and long intervals. We assessed multiple exploratory bifactor models, each with a different number of specific orthogonal factors alongside the general factor. The bifactor model with a specific factor F1 and general factor G was selected based on the best goodness of fit relative to both the unidimensional model and a bifactor model with two specific factors, and due to its interpretability advantage over the correlated-factors solution (Table 1). Additional bifactor indices in Table 1 attributed 79 % and 21 % of the common variance to F1 and G factors, respectively, with the former taking up the majority of reliable explained variance (relative omega ω and PUC). Similarly, the IECVs (Fig. 3e–f) indicate that the mid-range stimuli (1000–2000 ms) load principally on F1 (>80 %). Both the F1 and G factors exhibited good construct reliability with H = 0.96 and



	G loadings	F1 loadings	IECV
400	.61	.38	.28
600	.59	.62	.52
800	.57	.67	.58
1000	.37	.74	.80
1200	.14	.90	.98
1400	.00	.89	1.00
1600	-.20	.89	.95
1800	-.30	.87	.90
2000	-.32	.86	.88
2200	-.42	.81	.79
2400	-.44	.77	.76

Fig. 3. Factor extraction and latent variable models of reproduction times (Experiment 1, $N = 302$) a. The correlation matrix for reproduced durations across stimulus intervals. b. Observed and simulated eigenvalues of parallel analysis indicate a two-factor solution. c. Three-factor EGA solution. Edges (lines) are partial correlations adjusting for all other stimulus intervals in the network with edge thickness corresponding to magnitude and edge colour corresponding to direction (green: positive; red: negative); nodes (markers) reflect stimulus intervals and are colour-coded by the EFA factor. d. Factor loadings for each stimulus interval for correlated-factors and bifactor models. Interval cross-loadings (both factors with loadings larger than 0.3) for the correlated-factors model are displayed within the vertical grey lines. The grey shaded area highlights the region with low factor loadings (−0.3 to 0.3). e–f. Factor loadings for the bifactor model with a general factor (G), single specific factor (F1), and standardised loadings for each stimulus interval. Line thickness increases for greater absolute loadings, line colour denotes loading direction (green: positive; red: negative), and pale colours denote small factor loadings (< absolute 0.3). IECV = item explained common variance. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

0.72, respectively.

The factor loadings for the bifactor model (Fig. 3e–f) clearly indicate a pattern that aligns with the Bayesian central tendency effect (Acerbi, Wolpert, and Vijayakumar, 2012; Lejeune and Wearden, 2009; Sadibolova and Terhune, 2022). Specifically, the loadings show one factor (F1) on which all reproduction times load, reflecting a general timing function. In contrast, the second factor (G) includes only loadings (>|0.30|) corresponding to the most extreme (i.e., shortest and longest) intervals; notably, the direction of these loadings flipped at the interval midrange (1200-1600 ms), suggesting a process differentially recruited at these extremes. A plausible interpretation of G is that it reflects the Bayesian central tendency process, with short and long intervals weighted toward the mean of the stimulus intervals range. The magnitude of the central tendency bias is typically indicated by a slope coefficient in a model with each participant's reproduction times linearly regressed to stimulus intervals (Murai and Yotsumoto, 2016). To validate our interpretation, we conducted robust linear regression[1] analyses to examine whether factor scores predicted individual differences in central tendency bias, revealing that a larger bias (lower slopes) were associated with higher GF factor scores, $\beta = -0.15$, $t(300) = -24.27$, $p < .001$, $r = -0.81$, and lower F1 factor scores, $\beta = 0.11$, $t(300) = 14.20$, $p < .001$, $r = 0.58$. Notably, the factors were orthogonal, $r(300) < 0.01$, $SE < 0.01$, $p > .99$. Taken together, these results suggest that the latent structure of temporal reproduction performance comprises an interval-invariant general timing function and an ancillary function that subserves Bayesian prior weighting underlying the central tendency effect.

3.1.2. *Changepoint analysis*

Our subsequent objective was to investigate whether individual temporal reproduction performance encompasses changepoints that reflect the operation of distinct interval-specific timing functions. We evaluated three frequentist changepoint algorithms, all of which showed comparable results. In each case, the majority of participants' reproduction time patterns did *not* exhibit changepoints (85–97 %, Fig. 4a), with strong agreement between the scale- and trend-based changepoint models, $\chi^2(2, N = 302) = 46.55$, $p < .001$, $phi = 0.39$ [CI: 0.29, 0.51], and between the mean- and trend-based models, $\chi^2(2, N = 302) = 10.44$, $p = .028$, $phi = 0.19$ [CI: 0.09, 0.30]. In contrast, the odds of identifying scale-based changepoints were not significantly associated with those for the mean-based changepoints, $\chi^2(2, N = 302) = 0.04$, $p = .56$, $phi = 0.01$ [CI: 0, 0.12]. Put differently, the inflection points marking changes in variance did not correspond with those for changes in cumulative

[1] The models explained significant amount of variance in slopes, $F(2,300) = 589$, $p < .001$, $R^2_{adjusted} = 0.66$, $RMSE = 0.10$ (general factor predictor) and $F(2,300) = 205$, $p < .001$, $R^2_{adjusted} = 0.40$, $RMSE = 0.14$ (specific factor predictor).

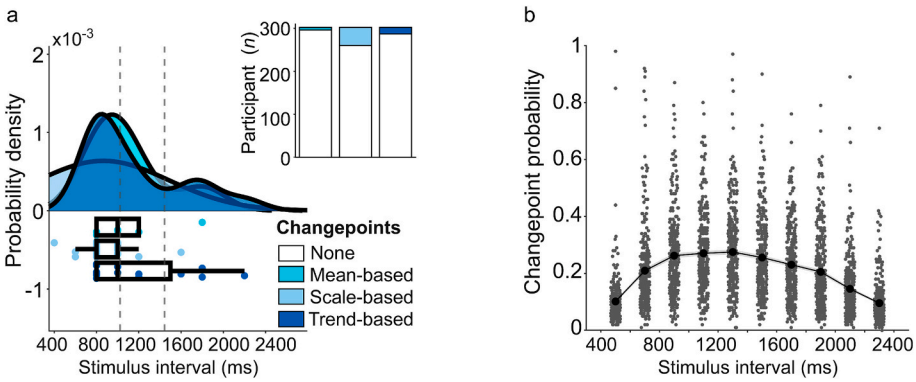


Fig. 4. Changepoint analysis applied to reproduction times in a visual temporal reproduction task in Experiment 1 ($N = 302$). a. Probability density functions and boxplots for the interval location of changepoints in reproduction times detected with the scale-, trend-, and mean-based frequentist models. The vertical dashed lines denote the cross-loadings in the EFA. Each coloured and white bar-proportion ratio in the inset is a representation of detected and not detected changepoints in the data, respectively. b. Bayesian implementation of changepoint analysis based on the cumulative mean. The plot shows probabilities of a changepoint for each participant (small markers), and across participants (mean [larger markers] and standard error [shaded area]).

mean. The detected changepoints tended to peak around 1000 ms, near to the boundary of the region with high cross-loadings in the EFA (Figs. 3d and 4). Fig. 4a shows that although the scale-based changepoint detection yielded more changepoints (15 % of participants), it also showed higher heterogeneity. Lastly, the Bayesian approach to changepoint analysis produced comparable findings, with <30 % probability of a changepoint for all stimulus intervals with a slightly increased probability in the 900-1300 ms interval range, which again overlaps with the region of high cross-loadings in the EFA (Fig. 4b). Cumulatively, these results do not provide any reliable evidence for an interval changepoint in reproduction times.

3.1.3. Synthesis of analytical approaches

The failure to observe robust changepoints across a large proportion of participants aligns with the bifactor analysis results. Logistic regression analyses showed that scale-based changepoint detection decreased 0.4 times with increasing GF factor scores, $\beta = -0.91$, $t(300) = 5.04$, $SE = 0.18$, $p < .001$, but was non-significantly related to F1 factor scores, $\beta = -0.05$, $t(300) = 0.29$, $SE = 0.17$, $p = .77$. Higher GF factor scores were also significant negative predictors of mean-based and trend-based changepoint detection, $\beta = -1.39$, $t(300) = 3.40$, $SE = 0.41$, $p < .001$, and $\beta = -1.61$, $t(300) = 4.99$, $SE = 0.32$, $p < .001$, respectively. There were also trends for F1 scores to positively predict detection of mean-based changepoints, $\beta = 0.83$, $t(300) = 1.88$, $SE = 0.44$, $p = .058$, and trend-based changepoints, $\beta = 0.76$, $t(300) = 2.56$, $SE = 0.30$, $p = .009$. These findings indicate a relationship between participants' factor loadings and the detection of changepoints, suggesting a convergence between the analyses conducted at both the group and individual levels.

3.2. Experiment 2

Cumulatively, the findings from Experiment 1 suggest that variability in temporal reproduction performance is driven by two distinct systems: a general timing system supporting reproduction across intervals and a specific system that underlies the extent to which reproduction is weighted or shaped by temporal context (Acerbi, Wolpert, and Vijayakumar, 2012; Lejeune and Wearden, 2009; Sadibolova and Terhune, 2022). Our second experiment sought to replicate and circumvent potential confounds in Experiment 1. If the interpretation of two correlated factors reflecting two overlapping timing systems holds, restricting the interval range, and specifically removing subsecond intervals, would be expected to yield a unidimensional factorial solution. By contrast, if our Bayesian interpretation of the bifactor model results is correct, we would expect to observe a similar solution even after restricting the interval range. We further opted to include a denser interval spacing in order to yield a larger stimulus interval pool. Finally, the stimulus intervals were arranged on a logarithmic, rather than linear, scale to address the smaller central tendency bias for short intervals and increase the discernibility of long intervals.

Upon first inspection, the reproduction times in Experiment 2 (Fig. 5) bore a striking resemblance with those in Experiment 1 (Fig. 2), with performance following the increase in stimulus intervals before levelling off and yielding a larger temporal bias for long intervals.

3.2.1. Dimensionality assessment

The correlations among reproduction times (Fig. 6a) were higher than in Experiment 1 (Fig. 3a), plausibly because of the inclusion of densely spaced stimulus intervals. Bartlett's test (Bartlett, 1951) was significant, $\chi^2(51) = 5673.32$, $p < .001$, and the KMO was 0.97, thereby confirming the suitability of the data for factor analysis. Parallel analysis, exploratory graph analysis, and the MAP statistics all indicated a two-factor solution (Fig. 6b–c; Supplemental Materials). In the correlated-factors model (Fig. 6d), short intervals primarily loaded on one factor whereas long intervals primarily loaded on the second factor. Although this pattern closely aligns with that observed in Experiment 1, the factor covariance was notably larger, $r(300) = 0.71$, $SE = 0.04$, $p < .001$. The fit statistics (Table 2) confirmed superior fit for the correlated-factors model over the unidimensional model, thus replicating the findings from Experiment 1.

To address the inherent interpretability issue in the correlated-factors model (see Experiment 1), we assessed the fit of different bifactor models with varying numbers of specific factors. The final model included a general factor and a specific factor (F1) (Table 2), based on the best goodness of fit relative to both the unidimensional model and a bifactor model with two specific factors, and due to its interpretability advantage over the correlated-factors model (see Experiment 1). Nevertheless, the bifactor indices indicate that the unidimensional model could still be a viable solution despite its worse model fit statistics (Table 2).

As can be seen in Fig. 6d–f, reproduction times for all stimulus intervals had high loadings on the general factor whereas only the most extreme intervals loaded (in opposite directions) on F1. In a conceptual replication of Experiment 1, the latter loadings flipped in the mid-interval range such that this factor was characterised by moderate positive loadings for the three shortest intervals and moderate negative loadings for the two longest intervals, reflecting a pattern consistent with the central tendency bias. In further support of this interpretation, as in Experiment 1, the steeper slopes for reproduction times regressed on stimulus intervals (reflecting a greater bias toward the mean of a stimulus interval set) were associated with lower F1 factor scores, $\beta =$

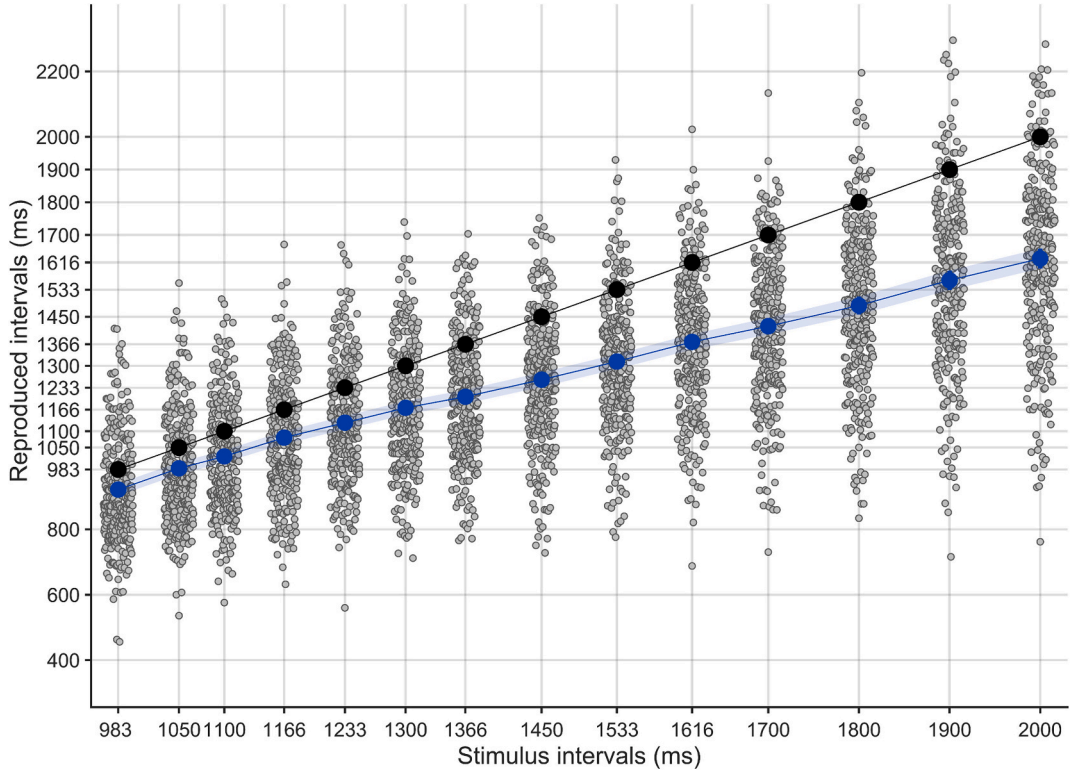


Fig. 5. Reproduced intervals in a visual temporal reproduction task in Experiment 2 (N = 302). Grey circles represent median reproduced intervals for each participant and stimulus interval. Blue circles and the blue shaded area represent sample medians and 95 % confidence intervals, respectively. Black circles denote veridical stimulus intervals and are included for reference. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

−0.21, $t(300) = 3.73$, $p < .001$, $r = -0.20$ and higher general factor scores, $\beta = 0.19$, $t(300) = 3.59$, $p < .001$, $r = 0.20$ (the factor scores did not significantly correlate, $r(300) = 0.01$, $p = .82$). Taken together, these results closely replicate those of Experiment 1 and suggest that the latent structure of temporal reproduction comprises a general timing factor and a second factor reflecting the central tendency bias.

3.2.2. Changepoint analyses

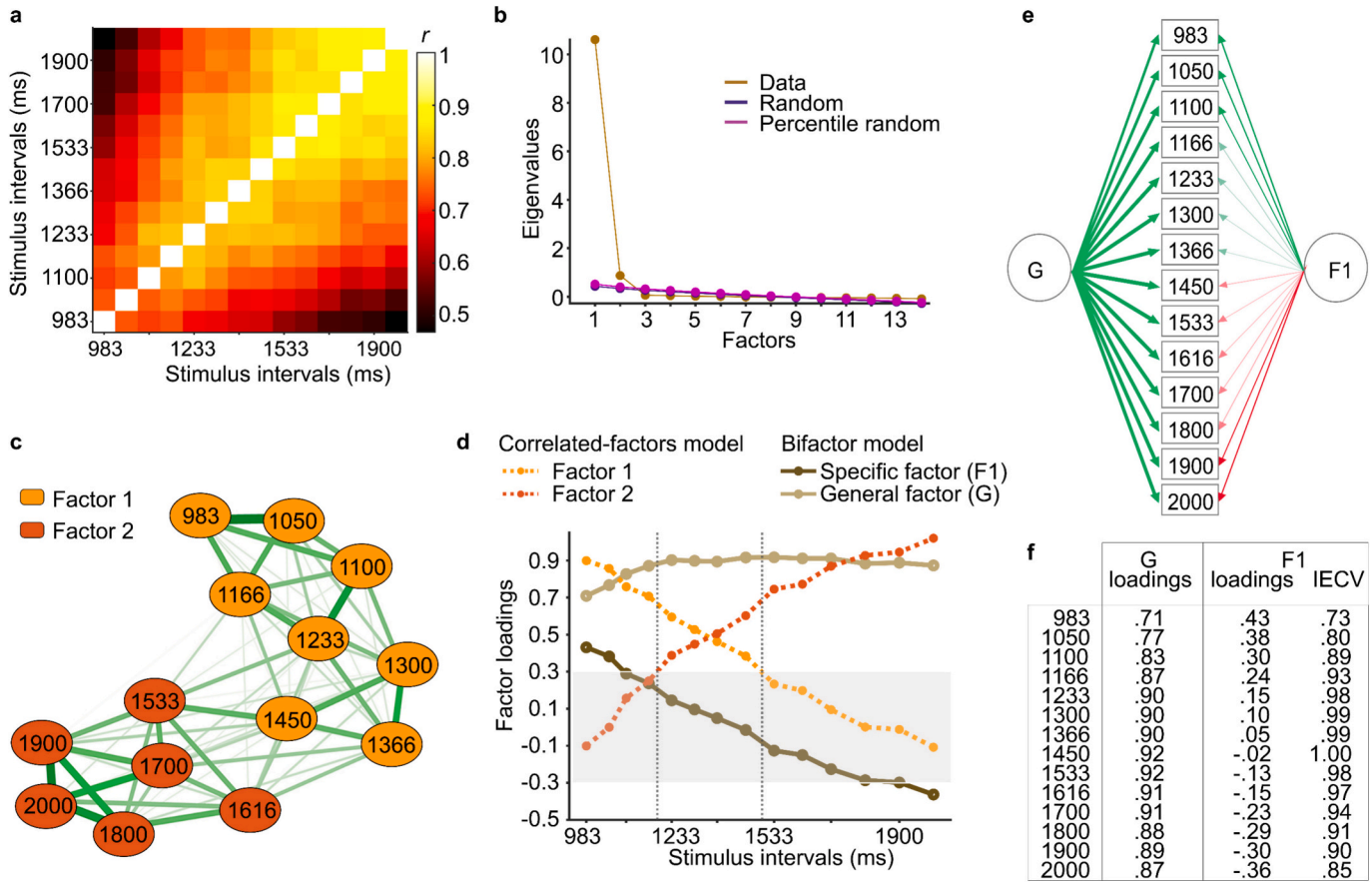
Although broadly in agreement, the results of the changepoint analyses deviated somewhat from those of Experiment 1. The scale-based (10.60 %) and trend-based (18.51 %) prevalence of identified changepoints in the sample was low, albeit higher than in Experiment 1 (Fig. 7a). Although we replicated the Fisher's exact test of association for changepoint detection across these two methods, $\chi^2(2, N = 302) = 12.07$, $p < .011$, $phi = 0.20$ [CI: 0.10, 0.31], changepoints exhibited flat probability density functions in contrast to the peaked distributions in Experiment 1 (Fig. 4a). Larger changepoint variability may be a consequence of a narrower stimulus range assessed on a finer scale. By contrast, the rate of the *mean*-based changepoint detection increased to 25.83 % and the peak of the probability density function for the detected changepoints migrated closer to the arithmetic mean of the stimulus intervals (Fig. 7a). Changepoint detection with this method did not significantly relate to that derived via the scale-based and trend-based detection methods, $\chi^2(2, N = 302) = 0.10$, $p = .83$, $phi = 0.02$ [CI: 0, 0.14] and $\chi^2(2, N = 302) = 2.05$, $p = .18$, $phi = -0.08$ [CI: −0.20, 0.0], respectively. Moreover, the mean-based frequentist and Bayesian changepoint analyses did not align (Fig. 7b) even though they both search for cumulative mean changes. We therefore exercise caution in interpreting these results and posit that the frequentist mean-based modelling may have picked up on inflections in log-spaced stimulus interval space specific to the included stimulus interval set. Notwithstanding the discrepancy in mean-based changepoint analysis, we again found little evidence consistent with the possibility that reproduction times included reliable inflection points suggestive of interval-specific timing systems.

3.2.3. Synthesis of analytical approaches

The convergence between the results of the bifactor model and changepoint analyses was replicated: higher F1 factor scores, plausibly reflecting the central tendency bias, were negatively associated with scale- and trend-based changepoint detection, $\beta = -0.60$, $t(300) = 3.00$, $SE = 0.20$, and $\beta = -0.62$, $t(300) = 3.73$, $SE = 0.17$, $p < .001$, respectively. By contrast, changepoint detection with both algorithms was not significantly related to bifactor general factor scores, $\beta = 0.22$, $t(300) = 1.20$, $SE = 0.18$, $p = .23$ and $\beta = 0.22$, $t(300) = 1.50$, $SE = 0.15$, $p = .13$, respectively. The mean-based changepoint detection was not predicted by bifactor model factor scores, F1: $\beta = -0.16$, $t(300) = 1.02$, $SE = 0.14$, $p = .31$ and general factor: $\beta = -0.06$, $t(300) = 0.43$, $SE = 0.13$, $p = .67$, showing no convergence across two analytical approaches.

3.3. General discussion

The objective of this study was to characterize the latent structure of temporal reproduction and to evaluate competing accounts of this structure. We were motivated to determine whether this latent structure would reflect the presence of distinct or overlapping mechanisms suggestive of different interval-specific timing systems, a unidimensional system, or multiple non-interval specific systems. Toward this end, in two large experiments, we applied multiple exploratory factor analysis and structural equation modelling approaches to reproduction times and further sought to corroborate these findings using multiple changepoint algorithms. The results of both experiments converged on a similar latent structure comprising two distinct factors that were not interval range-specific. The first, dominant, factor seemed to reflect a general



	G loadings	F1 loadings	IECV
983	.71	.43	.73
1050	.77	.38	.80
1100	.83	.30	.89
1166	.87	.24	.93
1233	.90	.15	.98
1300	.90	.10	.99
1366	.90	.05	.99
1450	.92	-.02	1.00
1533	.92	-.13	.98
1616	.91	-.15	.97
1700	.91	-.23	.94
1800	.88	-.29	.91
1900	.89	-.30	.90
2000	.87	-.36	.85

Fig. 6. Factor extraction and latent variable models of reproduction times (Experiment 2, $N = 302$) a. The correlation matrix for reproduced times across stimulus intervals. b. Observed and simulated eigenvalues of parallel analysis and c. Two-factor EGA solution. Edges (lines) are partial correlations adjusting for all other stimulus intervals in the network with edge thickness corresponding to magnitude and edge colour corresponding to direction (green: positive; red: negative); nodes (markers) reflect stimulus intervals and are colour-coded by the EFA factor. d. Factor loadings for correlated-factors and bifactor models. Interval cross-loadings (both factors with loadings larger than 0.3) are displayed within the vertical grey lines. The grey shaded area highlights the region with low factor loadings (−0.3 to 0.3). e–f. Factor loadings for the bifactor model with a general factor (G), single specific factor (F1). Line thickness increases for greater absolute loadings, line colour denotes loading direction (green: positive; red: negative), and pale colours denote small factor loadings (<absolute 0.3). IECV = item explained common variance. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

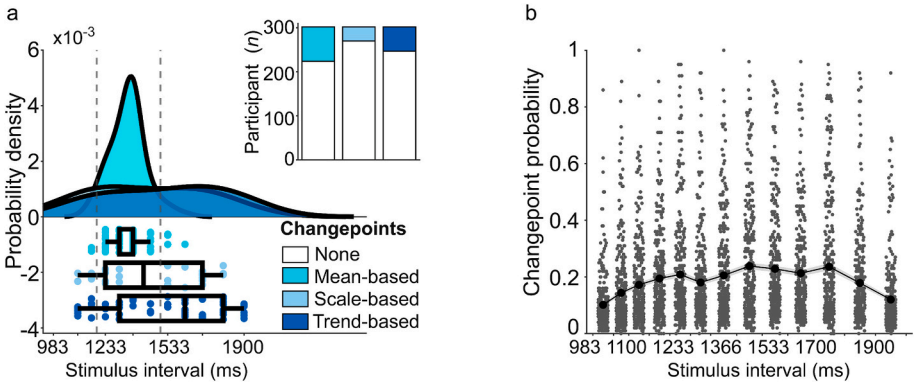


Fig. 7. Changepoint analysis for Experiment 2 ($N = 302$). a. Probability density functions for changepoints in reproduced intervals detected with the mean-, scale-, and trend-based frequentist models. The vertical dashed lines denote the cross-loadings in EFA. Each coloured and white bar-proportion ratio in the inset is a representation of detected and not detected changepoints in the data, respectively. b. Bayesian implementation of changepoint analysis based on the cumulative mean. The plot shows probabilities of a changepoint for each participant (small markers) and across participants (mean [larger markers] and standard error [shaded area]).

timing system that was recruited across intervals whereas the second factor appears to reflect the central tendency bias. Changepoint analyses provided further indirect evidence for this formulation. These results align with Bayesian models of time perception (Acerbi, Wolpert, and Vijayakumar, 2012; Jazayeri and Shadlen, 2010; Sadibolova and Terhune, 2022) and highlight the value of latent variable modelling in dissociating different cognitive processes underlying human time perception.

Notably, our initial factor analysis appeared to support a different conclusion. We replicated observations from previous research (two correlated factors) that would typically be interpreted as evidence for two overlapping timing systems subserving specific interval ranges (Rammsayer and Troche, 2014). Our data would suggest that in the ~1–1.5 s range, the system for short intervals fades as the system for long intervals becomes more dominant with increasing stimulus duration. This is consistent with prior work identifying 1 s (Buhusi and Meck, 2005; Gooch, Wiener, Hamilton, and Coslett, 2011), 1.2 s (Grondin, Ouellet, and Roussel, 2004), or 1.5 s (Gibbon, Malapani, Dale, and Gallistel, 1997; Grondin, 2012) as boundaries between these timing systems. However, this result contrasts with studies suggesting that the boundary lies either at the lower end of our interval range (~100–500 ms) or much later, at ~3–4 s (Lewis and Miall, 2009; Rammsayer and Pichelmann, 2018). If patterns in reproduction times were indeed driven by two timing systems with interval-specific roles, we would expect to observe reliable changepoints in a significant proportion of participants, particularly within the interval ranges where these correlated factors overlapped. However, changepoints indicating a shift from one system to another were found in only a small minority of participants and did not reliably align with this overlapping interval region.

This discrepancy between our factor analysis and changepoint findings was clarified when we sought to specify the interval subranges subserved uniquely by each timing system. Instead of identifying distinct systems in either experiment, our modelling revealed two interval-invariant and independent timing systems. One appeared to represent a general timing process active across all stimulus intervals, and the other seemed to reflect a system specifically involved at the extremes of the investigated interval range. We interpret this latter process as aligning with contemporary Bayesian accounts (Acerbi, Wolpert, and Vijayakumar, 2012; Jazayeri and Shadlen, 2010), reflecting a central tendency effect in which the reproduction of both short and long intervals shifts toward the mean of the stimulus interval range. This interpretation was further supported by the observation that the central tendency factor correlated with within-participant regression slopes, which provide an estimate of central tendency bias (Acerbi, Wolpert, and Vijayakumar, 2012; Cicchini, Arrighi, Cecchetti, Giusti, and Burr, 2012), and predicted the likelihood of detecting changepoints in individual participants' reproduced intervals.

The present results have implications for the measurement and interpretation of temporal reproduction performance in a variety of contexts. This work demonstrates that applying advances in factor analytic modelling (bifactor models; Markon, 2019) to interval timing response patterns can yield results that complement those derived through computational modelling (Acerbi, Wolpert, and Vijayakumar, 2012; Jazayeri and Shadlen, 2010; Petzschner, Glasauer, and Stephan, 2015). Accordingly, greater empirical attention to the overlap and divergences between these approaches is warranted. Relatedly, alongside Bayesian accounts of temporal performance, the present results highlight the importance of dissociating the latent cognitive processes subserving reproduction times from those supporting the precision-weighting of temporal priors underlying the central tendency bias (Cassidy et al., 2018). A considerable amount of research using temporal reproduction tasks conflates these processes, thereby hindering progress in understanding their neurophysiological and cognitive correlates. This work thus may provide an analytic approach that could be harnessed in computational factor modelling, whereby factor analytic and computational modelling approaches are integrated in attempts to characterize symptom heterogeneity in clinical and non-clinical samples (Wise, Robinson, and Gillan, 2023). In particular, we expect that this method could inform future research on distortions in time perception in response to pharmacological agents (Sadibolova et al., 2023; Wittmann et al., 2007; Yanakieva et al., 2019) and in psychiatric and neurological disorders (Bschor et al., 2004; DiMarco et al., 2023; Thoenes and Oberfeld, 2017).

Despite the advances afforded by this work, interpretation of this research should be constrained by the limitations of our experiments. Although our data strongly suggest that most of the variance in temporal reproduction is attributable to an interval-invariant cognitive function, or set of functions, the data do not lend specific support to any of the competing accounts of this function (Rammsayer and Pichelmann, 2018; Wiener, Matell, and Coslett, 2011). In addition, although we maintain that the present results are at odds with the operation of two interval-specific timing systems, it needs to be acknowledged that this only holds for reproduction times in the studied interval range. Our results provide a compelling account of the latent structure of temporal reproduction in this interval range. That is, although these systems were not observed in our data, we cannot rule out the possibility that they would be identified with other psychophysical interval timing tasks or extended interval ranges or with neurophysiological measures. Our data merely suggest that distinct overlapping systems do not provide the most empirically-grounded interpretation of variability in temporal reproduction in this interval range.

Taken together, these experiments found that the latent structure of temporal reproduction is best modelled by a general interval-invariant timing system and an ancillary system underlying the central tendency bias that plausibly reflects the precision weighting of temporal priors within Bayesian accounts of time perception (Acerbi, Wolpert, and Vijayakumar, 2012; Cicchini, Arrighi, Cecchetti, Giusti, and Burr, 2012; Jazayeri and Shadlen, 2010; Sadibolova and Terhune, 2022). Multiple lines of evidence converge to demonstrate that this model provides a better explanation of variability in temporal reproduction than a model attributing this variability to differential interval-specific timing systems. This work helps to empirically and theoretically integrate factor analytic and computational modelling approaches in the study of time perception and has implications for the measurement and interpretation of interval timing performance in a range of contexts.

CRediT authorship contribution statement

Renata Sadibolova: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Project administration, Formal analysis, Data curation. **Curtis Widmer:** Writing – review & editing, Validation, Methodology, Investigation, Conceptualization. **Zoe Fletcher:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Soraya Weill:** Writing – review & editing, Software, Methodology, Conceptualization. **Devin Terhune:** Writing – review & editing, Validation, Supervision, Software, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by the Biotechnology and Biological Sciences Research Council grant BB/R01583X/1 and by the Research and Enterprise Committee Fund (Goldsmiths, University of London) to DBT.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2025.106078.

Data availability

The data is available on the OSF website. The link is included in the

manuscript.

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), Article e1002771. https://doi.org/10.1371/journal.pcbi.1002771

Bangert, A. S., Reuter-Lorenz, P. A., & Seidler, R. D. (2011). Dissecting the clock: Understanding the mechanisms of timing across tasks and temporal intervals. *Acta Psychologica, 136*(1), 20–34. https://doi.org/10.1016/j.actpsy.2010.09.006

Barry, D., & Hartigan, J. A. (1993). A Bayesian analysis for change point problems. *Journal of the American Statistical Association, 88*(421), 309. https://doi.org/10.2307/2290726

Bonifay, W. E., Reise, S. P., Scheines, R., & Meijer, R. R. (2015). When are multidimensional data unidimensional enough for structural equation modeling? An evaluation of the DETECT multidimensionality index. *Structural Equation Modeling: A Multidisciplinary Journal, 22*(4), 504–516. https://doi.org/10.1080/10705511.2014.938596

Brody, C. D. (2003). Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. *Cerebral Cortex, 13*(11), 1196–1207. https://doi.org/10.1093/cercor/bhg100

Bschor, T., Ising, M., Bauer, M., Lewitzka, U., Skerstupeit, M., Muller-Oerlinghausen, B., & Baethge, C. (2004). Time experience and time judgment in major depression, mania and healthy subjects. A controlled study of 93 subjects. *Acta Psychiatrica Scandinavica, 109*(3), 222–229. https://doi.org/10.1046/j.0001-690X.2003.00244.x

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews. Neuroscience, 6*(10), 755–765. https://doi.org/10.1038/nrn1764

Buonomano, D. V., Bramen, J., & Khodadadifar, M. (2009). Influence of the interstimulus interval on temporal processing and learning: Testing the state-dependent network model. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1525), 1865–1873. https://doi.org/10.1098/rstb.2009.0019

Cassidy, C. M., Balsam, P. D., Weinstein, J. J., Rosengard, R. J., Slifstein, M., Daw, N. D., … Horga, G. (2018). A perceptual inference mechanism for hallucinations linked to striatal dopamine. *Current Biology, 28*(4), 503–514.e4. https://doi.org/10.1016/j.cub.2017.12.059

Chmielewski, M., & Kucker, S. C. (2020). An MTurk crisis? Shifts in data quality and the impact on study results. *Social Psychological and Personality Science, 11*(4), 464–473. https://doi.org/10.1177/1948550619875149

Cicchini, G. M., Arrighi, R., Cecchetti, L., Giusti, M., & Burr, D. C. (2012). Optimal encoding of interval timing in expert percussionists. *The Journal of Neuroscience, 32*(3), 1056–1060. https://doi.org/10.1523/JNEUROSCI.3411-11.2012

Dehling, H., Fried, R., & Wendler, M. (2020). A robust method for shift detection in time series. *Biometrika, 107*(3), 647–660. https://doi.org/10.1093/biomet/asaa004

DiMarco, E. K., Sadibolova, R., Jiang, A., Liebenow, B., Jones, R. E., Haq, I.u., … Kishida, K. T. (2023). Time perception reflects individual differences in motor and non-motor symptoms of Parkinson's disease. *Parkinsonism & Related Disorders, 114*, Article 105800. https://doi.org/10.1016/j.parkreldis.2023.105800

Edwards, C. J., Alder, T. B., & Rose, G. J. (2002). Auditory midbrain neurons that count. *Nature Neuroscience, 5*(10), 934–936. https://doi.org/10.1038/nn916

Erdman, C., & Emerson, J. W. (2007). bcp: An R package for performing a Bayesian analysis of change point problems. *Journal of Statistical Software, 23*(3). https://doi.org/10.18637/jss.v023.i03

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*(3), 279–325. https://doi.org/10.1037/0033-295X.84.3.279

Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. R. (1997). Toward a neurobiology of temporal cognition: Advances and challenges. *Current Opinion in Neurobiology, 7*(2), 170–184. https://doi.org/10.1016/S0959-4388(97)80005-0

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLoS One, 12*(6), Article e0174035. https://doi.org/10.1371/journal.pone.0174035

Gooch, C. M., Wiener, M., Hamilton, A. C., & Coslett, H. B. (2011). Temporal discrimination of sub- and suprasecond time intervals: A voxel-based lesion mapping analysis. *Frontiers in Integrative Neuroscience, 5*. https://doi.org/10.3389/fnint.2011.00059

Gorsuch, R. L. (2013). *Factor analysis*. Lawrence Erlbaum Associates, Inc., Publishers. Psychology Press. https://doi.org/10.4324/9780203781098

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics, 72*(3), 561–582. https://doi.org/10.3758/APP.72.3.561

Grondin, S. (2012). Violation of the scalar property for time perception between 1 and 2 seconds: Evidence from interval discrimination, reproduction, and categorization. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 880–890. https://doi.org/10.1037/a0027188

Grondin, S., Meilleur-Wells, G., & Lachance, R. (1999). When to start explicit counting in a time-intervals discrimination task: A critical point in the timing process of humans. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 993–1004. https://doi.org/10.1037/0096-1523.25.4.993

Grondin, S., Ouellet, B., & Roussel, M.-E. (2004). Benefits and limits of explicit counting for discriminating temporal intervals. *Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale, 58*(1), 1–12. https://doi.org/10.1037/h0087436

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In *Structural Equation Modeling: Present and Future: A Festschrift in Honor of Karl Jöreskog* (pp. 195–216). Scientific Software International.

Hattori, M., Zhang, G., & Preacher, K. J. (2017). Multiple local solutions and geomin rotation. *Multivariate Behavioral Research, 52*(6), 720–731. https://doi.org/10.1080/00273171.2017.1361312

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2), 179–185. https://doi.org/10.1007/BF02289447

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience, 13*(8), 1020–1026. https://doi.org/10.1038/nn.2590

Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika, 76*(4), 537–549. https://doi.org/10.1007/s11336-011-9218-4

Kline, R. B. (2023). *Principles and Practice of Structural Equation Modeling - Fifth Edition.* Guilford Press.

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. https://doi.org/10.1080/09541440802453006

Lewis, P., & Miall, R. (2003). Brain activation patterns during measurement of sub- and supra-second intervals. *Neuropsychologia, 41*(12), 1583–1592. https://doi.org/10.1016/S0028-3932(03)00118-0

Lewis, P. A., & Miall, R. C. (2009). The precision of temporal judgement: Milliseconds, many minutes, and beyond. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1525), 1897–1905. https://doi.org/10.1098/rstb.2009.0020

Lusk, N. A., Petter, E. A., & Meck, W. H. (2020). A systematic exploration of temporal bisection models across sub- and supra-second duration ranges. *Journal of Mathematical Psychology, 94*, Article 102311. https://doi.org/10.1016/j.jmp.2019.102311

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika, 57*(3), 519–530. https://doi.org/10.1093/biomet/57.3.519

Markon, K. E. (2019). Bifactor and hierarchical models: Specification, inference, and interpretation. *Annual Review of Clinical Psychology, 15*(1), 51–69. https://doi.org/10.1146/annurev-clinpsy-050718-095522

Merchant, H., & Georgopoulos, A. P. (2006). Neurophysiology of perceptual and motor aspects of interception. *Journal of Neurophysiology, 95*(1), 1–13. https://doi.org/10.1152/jn.00422.2005

Merchant, H., Zarco, W., Bartolo, R., & Prado, L. (2008). The context of temporal processing is represented in the multidimensional relationships between timing tasks. *PLoS One, 3*(9), Article e3169. https://doi.org/10.1371/journal.pone.0003169

Mondok, C., & Wiener, M. (2023). Selectivity of timing: A meta-analysis of temporal processing in neuroimaging studies using activation likelihood estimation and reverse inference. *Frontiers in Human Neuroscience, 16*, Article 1000995. https://doi.org/10.3389/fnhum.2022.1000995

Murai, Y., & Yotsumoto, Y. (2016). Timescale- and sensory modality-dependency of the central tendency of time perception. *PLoS One, 11*(7), Article e0158921. https://doi.org/10.1371/journal.pone.0158921

Muthén, L. K., & Muthén, B. O. (2007). *Mplus User's Guide* (7th ed.). Muthén & Muthén.

Nani, A., Manuello, J., Liloia, D., Duca, S., Costa, T., & Cauda, F. (2019). The neural correlates of time: A meta-analysis of neuroimaging studies. *Journal of Cognitive Neuroscience, 31*(12), 1796–1826. https://doi.org/10.1162/jocn_a_01459

Osborne, J. W., & Costello, A. B. (2004). Sample size and subject to item ratio in principal components analysis. *Practical Assessment, Research, and Evaluation, 9*(1), 11. https://doi.org/10.7275/ktzq-jq66

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2021). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods, 54*(4), 1643–1662. https://doi.org/10.3758/s13428-021-01694-3

Petzschner, F. H., Glasauer, S., & Stephan, K. E. (2015). A Bayesian perspective on magnitude estimation. *Trends in Cognitive Sciences, 19*(5), 285–293. https://doi.org/10.1016/j.tics.2015.03.002

R Core Team. (2021). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing.

Rammsayer, T. H. (1993). On dopaminergic modulation of temporal information processing. *Biological Psychology, 36*(3), 209–222. https://doi.org/10.1016/0301-0511(93)90018-4

Rammsayer, T. H. (2008). Neuropharmacological approaches to human timing. In S. Grondin (Ed.), *Psychology of Time* (pp. 295–320). Emerald.

Rammsayer, T. H., & Altenmüller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception, 24*(1), 37–48. https://doi.org/10.1525/mp.2006.24.1.37

Rammsayer, T. H., & Brandler, S. (2004). Aspects of temporal information processing: A dimensional analysis. *Psychological Research Psychologische Forschung, 69*(1–2), 115–123. https://doi.org/10.1007/s00426-003-0164-3

Rammsayer, T. H., & Lima, S. D. (1991). Duration discrimination of filled and empty auditory intervals: Cognitive and perceptual factors. *Perception & Psychophysics, 50*(6), 565–574. https://doi.org/10.3758/BF03207541

Rammsayer, T. H., & Pichelmann, S. (2018). Visual-auditory differences in duration discrimination depend on modality-specific, sensory-automatic temporal processing: Converging evidence for the validity of the sensory-automatic timing hypothesis. *Quarterly Journal of Experimental Psychology, 71*(11), 2364–2377. https://doi.org/10.1177/1747021817741611

Rammsayer, T. H., & Troche, S. J. (2014). In search of the internal structure of the processes underlying interval timing in the sub-second and the second range: A confirmatory factor analysis approach. *Acta Psychologica, 147*, 68–74. https://doi.org/10.1016/j.actpsy.2013.05.004

Rammsayer, T. H., & Ulrich, R. (2001). Counting models of temporal discrimination. *Psychonomic Bulletin & Review, 8*(2), 270–277. https://doi.org/10.3758/BF03196161

Revelle, W., & Rocklin, T. (1979). Very simple structure: An alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research, 14*(4), 403–414. https://doi.org/10.1207/s15327906mbr1404_2

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment, 98*(3), 223–237. https://doi.org/10.1080/00223891.2015.1089249

Sadibolova, R., Murray-Lawson, C., Family, N., Williams, L. T. J., Luke, D. P., & Terhune, D. B. (2023). LSD microdosing attenuates the impact of temporal priors in time perception. *BioRxiv*. https://doi.org/10.1101/2023.04.14.536983

Sadibolova, R., Sun, S., & Terhune, D. B. (2021). Using adaptive psychophysics to identify the neural network reset time in subsecond interval timing. *Experimental Brain Research, 1*, 1–8. https://doi.org/10.1007/s00221-021-06227-0

Sadibolova, R., & Terhune, D. B. (2022). The temporal context in Bayesian models of interval timing: Recent advances and future directions. *Behavioral Neuroscience, 136*(5), 364–373. https://doi.org/10.1037/bne0000513

Schirmer, A. (2004). Timing speech: A review of lesion and neuroimaging findings. *Cognitive Brain Research, 21*(2), 269–287. https://doi.org/10.1016/j.cogbrainres.2004.04.003

Sohn, M.-H., & Carlson, R. A. (2003). Implicit temporal tuning of working memory strategy during cognitive skill acquisition. *The American Journal of Psychology, 116*(2), 239. https://doi.org/10.2307/1423579

Stauffer, C. C., Haldemann, J., Troche, S. J., & Rammsayer, T. H. (2012). Auditory and visual temporal sensitivity: Evidence for a hierarchical structure of modality-specific and modality-independent levels of temporal information processing. *Psychological Research, 76*(1), 20–31. https://doi.org/10.1007/s00426-011-0333-8

Stucky, B. D., & Edelen, M. O. (2015). Using hierarchical IRT models to create unidimensional measures from multidimensional data. In S. P. Reise, & D. A. Revicki (Eds.), *Handbook of Item Response Theory Modeling: Applications to Typical Performance Assessment* (pp. 183–206). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9781315736013.ch9.

Thoenes, S., & Oberfeld, D. (2017). Meta-analysis of time perception and temporal processing in schizophrenia: Differential effects on precision and accuracy. *Clinical Psychology Review, 54*, 44–64. https://doi.org/10.1016/j.cpr.2017.03.007

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41*(3), 321–327. https://doi.org/10.1007/BF02293557

Wang, J., & Wang, X. (2012). Structural equation modeling. In *Structural Equation Modeling: Applications Using Mplus*. John Wiley & Sons Ltd.. https://doi.org/10.1002/9781118356258

Weston, R., & Gore, P. A. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist, 34*(5), 719–751. https://doi.org/10.1177/0011000006286345

Wiener, M., Lohoff, F. W., & Coslett, H. B. (2011). Double dissociation of dopamine genes and timing in humans. *Journal of Cognitive Neuroscience, 23*(10), 2811–2821. https://doi.org/10.1162/jocn.2011.21626

Wiener, M., Matell, M. S., & Coslett, H. B. (2011). Multiple mechanisms for temporal processing. *Frontiers in Integrative Neuroscience, 5*(31). https://doi.org/10.3389/fnint.2011.00031

Wise, T., Robinson, O. J., & Gillan, C. M. (2023). Identifying transdiagnostic mechanisms in mental health using computational factor modeling. *Biological Psychiatry, 93*(8), 690–703. https://doi.org/10.1016/j.biopsych.2022.09.034

Wittmann, M., Carter, O., Hasler, F., Cahn, B. R., Grimberg, U., Spring, P., … Vollenweider, F. X. (2007). Effects of psilocybin on time perception and temporal control of behaviour in humans. *Journal of Psychopharmacology, 21*(1), 50–64. https://doi.org/10.1177/0269881106065859

Yanakieva, S., Polychroni, N., Family, N., Williams, L. T. J., Luke, D. P., & Terhune, D. B. (2019). The effects of microdose LSD on time perception: A randomised, double-blind, placebo-controlled trial. *Psychopharmacology, 236*(4), 1159–1170. https://doi.org/10.1007/s00213-018-5119-x

Zakay, D. (1990). The evasive art of subjective time measurement: Some methodological dilemmas. In R. A. Block (Ed.), *Cognitive Models of Psychological Time* (pp. 59–84). Lawrence Erlbaum Associates, Inc.