Journal of Experimental Psychology: General

Abstract Thinking Facilitates Aggregation of Information

Britt Hadar, Moshe Glickman, Yaacov Trope, Nira Liberman, and Marius Usher

Online First Publication, December 20, 2021. http://dx.doi.org/10.1037/xge0001126

CITATION

Hadar, B., Glickman, M., Trope, Y., Liberman, N., & Usher, M. (2021, December 20). Abstract Thinking Facilitates Aggregation of Information. *Journal of Experimental Psychology: General*. Advance online publication. http://dx.doi.org/10.1037/xge0001126

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

https://doi.org/10.1037/xge0001126

BRIEF REPORT

Abstract Thinking Facilitates Aggregation of Information

Britt Hadar[1, 2, 3], Moshe Glickman[4, 5], Yaacov Trope[6], Nira Liberman[1], and Marius Usher[1, 7]

[1] School of Psychological Sciences, Tel-Aviv University
[2] Department of Psychology, Princeton University
[3] Princeton School of Public and International Affairs, Princeton University
[4] Department of Experimental Psychology, University College London
[5] Max Planck University College London Centre for Computational Psychiatry and Ageing Research, London, United Kingdom
[6] Department of Psychology, New York University
[7] Sagol School of Neuroscience, Tel-Aviv University

Many situations in life (such as considering which stock to invest in, or which people to befriend) require averaging across series of values. Here, we examined predictions derived from construal level theory, and tested whether abstract compared with concrete thinking facilitates the process of aggregating values into a unified summary representation. In four experiments, participants were induced to think more abstractly (vs. concretely) and performed different variations of an averaging task with numerical values (Experiments 1–2 and 4), and emotional faces (Experiment 3). We found that the induction of abstract, compared with concrete thinking, improved aggregation accuracy (Experiments 1–3), but did not improve memory for specific items (Experiment 4). In particular, in concrete thinking, averaging was characterized by increased regression toward the mean and lower signal-to-noise ratio, compared with abstract thinking.

Keywords: construal level theory, abstraction, numerical averaging, emotion perception

Supplemental materials: https://doi.org/10.1037/xge0001126.supp

Imagine yourself in a new workplace, trying to figure out how friendly are your new coworkers. To form impressions about their dispositions, you sample their behavior (e.g., whether they greeted you in the elevator or invited you to lunch), and average those samples into a general value-representation (Anderson, 1981). Similarly, when watching products displayed in a shop (Yamanashi Leib et al., 2020), or a rapid stream of stock returns (Betsch et al., 2001; Vanunu et al., 2019), one might extract an overall value estimation to guide behavior. In all these cases, different exemplars that span across time and space are collapsed into one representative value (e.g., "I like a lot this coworker, this shop is expensive, or the stock is profitable"). Averaging is a common and intuitive way of extracting a representative value from a stream of exemplars (Anderson, 1971; Kahneman, 2011). Indeed, the rapid extraction of ensemble information has been documented in a wide array of ensemble properties ranging from simple features such as size (Ariely, 2001; Chong et al., 2008; Chong & Treisman, 2003; 2005), orientation (Dakin & Watt, 1997; Parkes et al., 2001), color (Maule et al., 2014), motion direction (Watamaniuk & McKee, 1998; Watamaniuk et al., 1989), and motion speed (Watamaniuk & Duchon, 1992); to more complex properties, such as gender (Haberman & Whitney, 2007), identity (de Fockert & Wolfenstein, 2009; Neumann et al., 2013), emotional facial expressions (Haberman et al., 2009), animacy (Yamanashi Leib et al., 2016), gaze direction (Florey et al., 2016; Sweeny & Whitney, 2014), attractiveness (Post et al., 2012; van Osch et al., 2015; Walker & Vul, 2014), traits (Asch, 1946; Eyal et al., 2011; Hamilton & Sherman, 1996), and even when deciding whether a basketball player's career earns him a place in the Hall of Fame (Brusovansky et al., 2019). For

Britt Hadar https://orcid.org/0000-0003-0863-8650

Britt Hadar and Moshe Glickman shared equal contribution. Nira Liberman and Marius Usher joined senior authorship.

Data is available at: https://osf.io/y4uxd/.

This research was supported by a United States-Israel Binational Science Foundation (BSF) Grant 2016090 to Nira Liberman and Yaacov Trope. Nira Liberman and Britt Hadar were supported by Grant 524/17 by the Israel Science Foundation (ISF). Marius Usher was supported by an ISF Grant 1413/17. The authors would like to thank Talya Rapp and Yasmeen Shamshoum for their help in running the experiments.

Correspondence concerning this article should be addressed to Britt Hadar, Department of Psychology, Princeton University, South Drive, Princeton, NJ 08540, United States, or Moshe Glickman, Department of Experimental Psychology, University College London, 26 Bedford Way, London, WC1H 0AP, United Kingdom. Email: britt.hadar@gmail.com or mosheglickman345@gmail.com

an extensive review of ensemble perception, the reader is referred to Whitney and Yamanashi Leib (2018).

Recent research has supported the idea that average extraction is automatic, or at least nonintentional (Betsch et al., 2001; Brusovansky et al., 2018; Khayat & Hochstein, 2018), and is based on a population-coding mechanism (Baek & Chong, 2020; Brezis et al., 2016; Brezis et al., 2018). Several factors, including set size (Brezis et al., 2015; Robitaille & Harris, 2011), and variance (Brezis et al., 2018; Solomon, 2010) were found to affect the accuracy of averaging. However, little is known about the impact of thinking-modes on the accuracy of intuitive (i.e., without explicit symbolic calculation) averaging. We specifically focus on concrete versus abstract thinking. Research conducted within construal level theory (CLT), shows that people form representations at varying degrees of abstraction (Gilead et al., 2020; Liberman & Trope, 2014; Trope & Liberman, 2010). A process of abstraction, in this view, occurs when distinctions between exemplars are disregarded, and a dimension of similarity is highlighted that places them into a category. For example, categorizing sandals and boots as footwear disregards "waterproof" and highlights "worn on feet." Averaging bears inherent relation to abstraction because when a stream of exemplars is represented by an average, differences between specific exemplars are disregarded, and their commonality (their origin in a common distribution) is emphasized.

We reasoned that because averaging is a process of abstraction wherein exemplars are integrated (with equal weights) into a single representative value, then abstract thinking (compared with concrete thinking) should enhance the accuracy of intuitive average estimation. We tested this prediction in three experiments (Experiments 1–3) that manipulated, within participants, abstract versus concrete thinking via the "why/how" mindset manipulation, which was originally developed by Freitas et al. (2004), and has been widely used since then. Indeed, a review by Burgoon et al. (2013) mentions 11 published articles that used this manipulation, and many more used it since then (Ding & Keh, 2017; Efrat-Treister et al., 2020; Gilead et al., 2014; Hadar et al., 2019; Hansen & Trope, 2013; Kille et al., 2017; Napier et al., 2018; Spunt et al., 2016; Stillman et al., 2017; Yudkin et al., 2020). Experiment 1 examined accuracy in a numerical averaging task (Brezis et al., 2015; Malmi & Samson, 1983; Rosenbaum et al., 2021), and Experiment 2 was a preregistered replication. Experiment 3 tested the effect of abstract thinking on averaging of emotional faces (which prevents the use of rule-based strategies). Experiment 4 was a control experiment that tested the effect of abstract mindset on memory for specific details. We predicted that abstract (compared with concrete) thinking would improve accuracy of average extraction (Experiments 1–3), but not memory for specific details (Experiment 4). We examined two mechanisms that may explain the effect of abstract thought: signal-to-noise ratio and temporal weighting biases.

Experiments

Mindset Manipulation

In all four experiments[1] we used the mindset induction manipulation developed by Freitas et al. (2004); see Supplemental Information (SI) for a detailed description of the manipulation. In the abstract mindset condition, participants were presented with four boxes, in which they answered four consecutive "why" questions, starting from, for example, "Why maintain good physical health". In the concrete mindset condition, participants answered four consecutive "how" questions, starting with the same behavior, for example, "How to maintain good physical health". (Figure 1A presents the manipulation).

Averaging Task

In all three experiments, following each mindset induction, the participants estimated the average of rapidly presented sequences of values (Brezis et al., 2015; Malmi & Samson, 1983). The estimation precision was measured via the Pearson correlation between the actual and the estimated averages across all trials (for each participant and mindset manipulation). Relying on correlation allows testing sensitivity to relative changes in the variable, while allowing for general biases.[2] In the General Discussion section, we also report analyses of actual deviations (i.e., root mean square deviation [RMSD]).

Experiment 1

Method

Participants

Fifty-six Tel-Aviv University students (35 women, M_{age} = 22.80, SD = 4.12) participated for a course credit. One participant left before finishing the task and was excluded from analysis.

Materials

In the numerical averaging task, each trial began with a central fixation cross (250 ms), after which a sequence of eight two-digit numbers (between 10 and 90) was presented at a rate of 2 Hz and participants were required to enter their estimation of the sequence average on a visual analog scale ranging from 0–100 (see Figure 1B, and SI for additional details). The numbers were drawn from a Gaussian distribution ($\mu = 50 + k$, $\sigma = 20$, $k \sim U[-15,15]$), with no successive repetitions. Participants completed 20 practice trials and 240 experimental trials.

Procedure

Participants first completed the practice trials of the numerical averaging task, then the mindset manipulation (either the concrete "how" or the abstract "why"). Then 60 trials of the numerical averaging task, such that two "why" blocks were completed consecutively and then two "how" blocks were completed consecutively by each participant. Whether a participants started with "why" or "how" blocks was counterbalanced between participants.

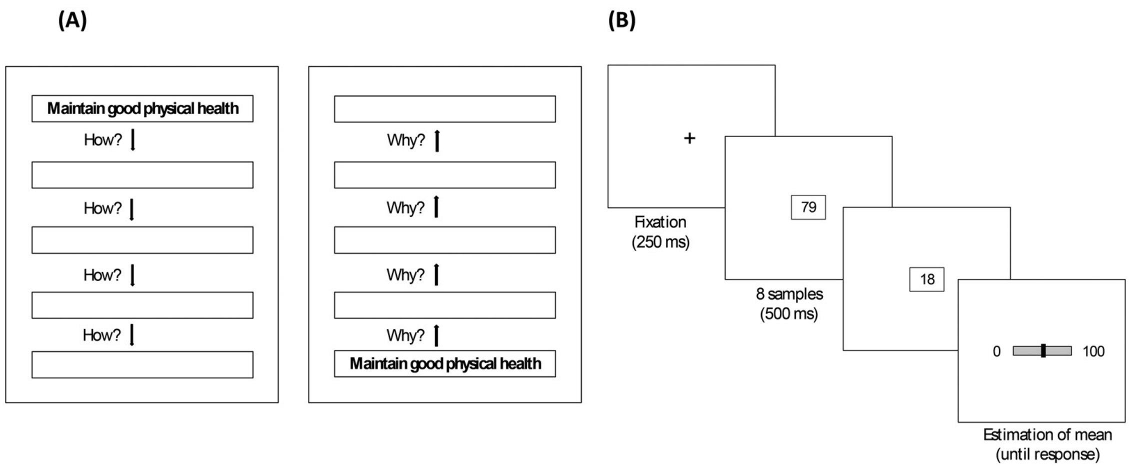
Results

Performance in the abstract and concrete thinking conditions was quantified by computing the Pearson correlation between participants' estimations and actual mean of each sequence (see Figure 2A–B for an illustration of two representative participants). Pearson

[1] Data is available at: https://osf.io/y4uxd/, see Hadar (2021).

[2] For example, if the average estimations are accurate but shifted by a constant, this would reduce the deviation score but preserve a high correlation between real and the estimated averages. Such a situation indicates a high discrimination sensitivity of the sequence-average.

Figure 1
Event Sequence in Experiments 1 and 2



Note. (A) Mindset induction manipulation: Participants complete the abstract mindset manipulation ("Why?") or the concrete mindset manipulation ("How?"). (B) Participants then observe a sequence of eight numbers and estimate their average on an analog scale (60 in each block). In Experiment 1 there were two blocks after a concrete mindset induction, two blocks after an abstract mindset induction. In Experiment 2 there was only one block after each manipulation.

correlations were transformed into Fisher's Z-scores (z = arctanh (r)) and submitted to 2 (mindset: abstract vs. concrete) × 2 (order: abstract block first vs. concrete block first) mixed-design ANOVA with mindset as a within-participants factor. The analysis revealed the predicted effect of mindset, whereby the Fisher-transformed correlations were significantly higher in the abstract mindset condition (M = .911, SD = .221) compared with the concrete mindset condition (M = .859, SD = .214), $F(1, 53) = 5.411$, $p = .024$, $\eta_p^2 = .093$. A main effect of order, $F(1, 53) = 8.210$, $p = .006$, $\eta_p^2 = .134$, indicated that correlations were higher for participants who completed the concrete mindset condition first compared to those who completed the abstract mindset condition first. The interaction was not significant, $F(1, 53) = .340$, $p = .562$, indicating that the effect of abstract thinking obtained irrespective of order (see Figure 2C for the group-level [z-transformed] Pearson correlations collapsed across order conditions). We confirmed that the difference between the abstract and concrete conditions remain significant if we use a nonparametric test to compare the raw correlations (i.e., without using Fisher z-transformation). To this end, we conducted a two-tailed paired-samples permutation test with 10,000 shuffles, which yielded a significant result ($p = .023$).

Experiment 2

Experiment 2 was a preregistered replication[3] of Experiment 1 with several changes. First, we cut the number of trials by half. Second, to generalize our finding across different underlying distributions, the values in Experiment 2 were sampled from a uniform distribution (in which all the numbers in the designated range have the same probability to be sampled). This makes the estimation more challenging by rendering inefficient heuristic strategies (e.g., basing the average on only two to three random samples.)

Method

Participants

Sixty-two Tel-Aviv University students (37 women, M_{age} = 23.69, SD = 3.27) participated in the experiment for course credit. Two participants were excluded; one was not a native Hebrew speaker, and one did not complete the manipulation forms.

Materials and Procedure

The procedure was the same as in Experiment 1, except that participants completed only one block of 60 trials under each mindset condition and the distribution from which numbers in each sequence were sampled was uniform ($U[45 + k, 55 + k]$, $k \sim U[-15, 15]$) rather than Gaussian.

Results

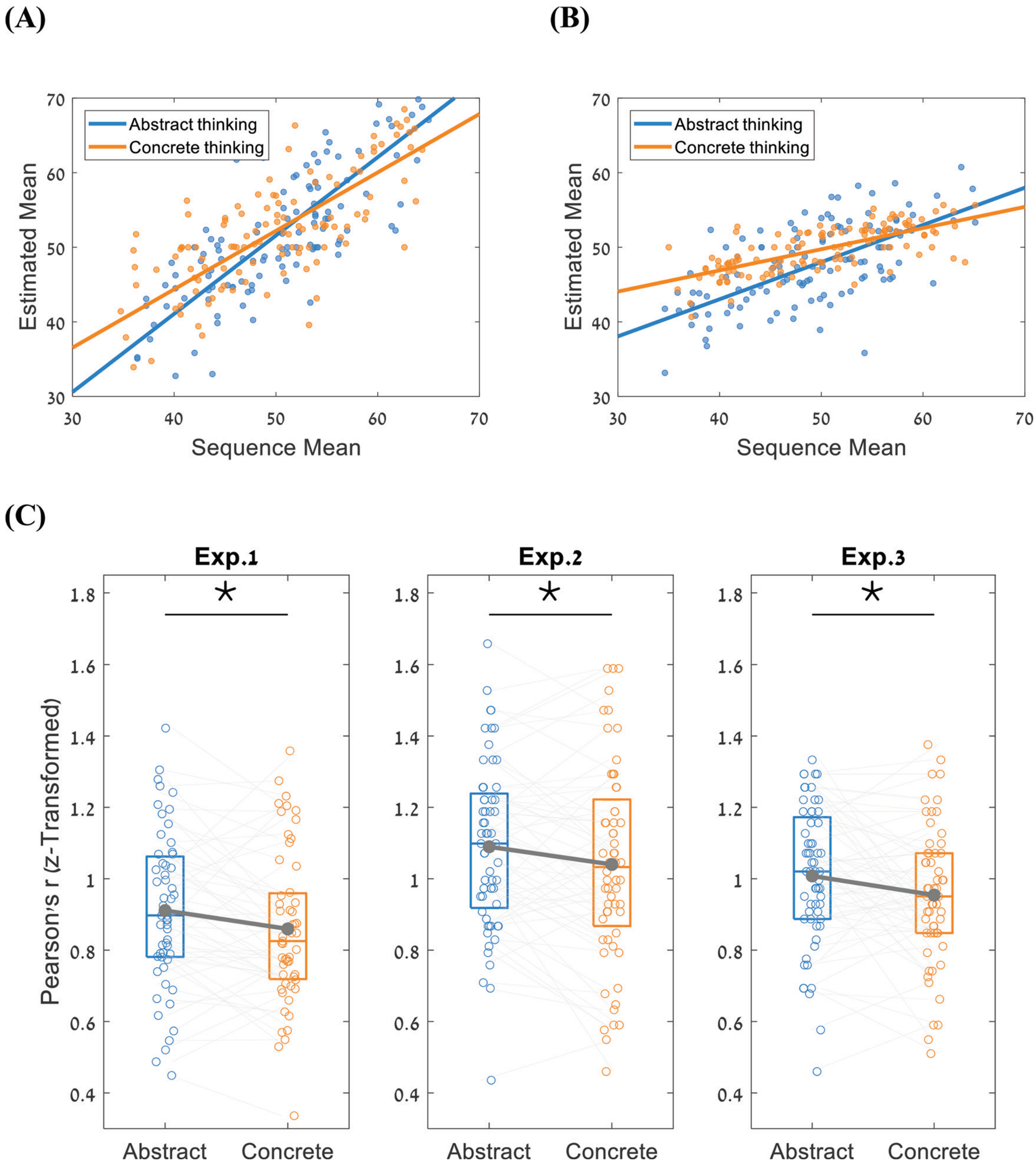
The same analysis as in Experiment 1 revealed the predicted effect of mindset, whereby Fisher-transformed correlations were higher in the abstract mindset condition (M = 1.091, SD = .234) compared with the concrete mindset condition, (M = 1.041, SD = .278), $F(1, 58) = 4.572$, $p = .037$, $\eta_p^2 = .073$. Neither order $F(1, 58) = .964$, $p = .330$, nor the interaction of mindset and order were significant, $F(1, 58) = .185$, $p = .669$ (see Figure 2C). The same pattern of results was obtained if we analyzed the data using a paired-sampled permutation test similar to that used in Experiment 1 ($p = .013$).

Experiment 3

The aim of Experiment 3 was to generalize the effect of abstraction on aggregation of information to more ecological and real-life

[3] http://aspredicted.org/blind.php?x=bd5245

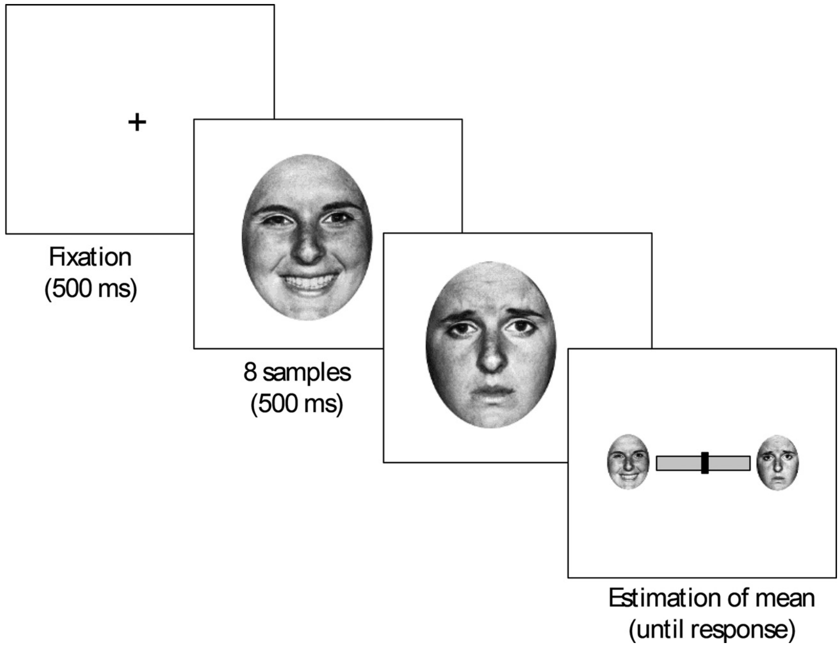
Figure 2
Results of Experiments 1–3



Note. (A–B) Correlations between the sequence actual mean and the estimated responses of two representative participants in the abstract (blue) and concrete (orange) thinking conditions. (C) Pearson correlations (z-transformed) by mindset condition (abstract vs. concrete) in Experiments 1–3. Blue and orange circles correspond to the correlations of each participant in the abstract and concrete conditions, respectively. Gray circles correspond to the conditions means. The central mark on each box-plot indicates the median, and the bottom and top edges indicate the 25th and 75th percentiles, respectively. $^*p < .05$. See the online article for the color version of this figure.

stimuli, which also prevents the use of rule-based computations. We chose extraction of the mean emotion from a set of emotional faces for several reasons. First, people may face a similar task on a daily basis when trying to understand their partners' or coworkers' feelings. Second, finding a common process that applies to both simple stimuli (numbers in Experiments 1 and 2) as well as to complex social stimuli (emotional faces) is not trivial and can shed light on a broader spectrum of instances in which abstraction plays a role.

Figure 3
Illustration of the Emotion Averaging Task in Experiment 3



Note. A fixation cross is presented for 500 ms, followed by a sequence of eight faces presented for 500 ms each. Participants then estimate the average emotional expression using an analog scale, presented until response. The authors adopted the morphed faces from Haberman et al. (2009).

Method

Participants

Sixty-two Tel-Aviv University students (38 women, M_{age} = 22.88, *SD* = 2.54) participated in the experiment for course credit. Two participants did not complete the manipulation forms and were excluded.

Materials

Mindset Manipulation. We used the same mindset manipulation as in Experiments 1 and 2.

Emotion Averaging Task. We adopted the morphed faces from Haberman et al. (2009). We used 50 different faces that were created by linearly interpolating between two emotion extremes of the same actor, taken from the Ekman gallery (Ekman & Friesen, 1976), such that each morphed face was one emotional unit sadder than the one before it. Thus, the values of the faces ranged between 1 (*happiest*) and 50 (*saddest*; see Figure 3). On each trial, faces were randomly sampled from a uniform distribution (with repetition).

Procedure

Participants were told that they will be presented with sequences of emotional faces (all of the same person, “Rachel”). Each sequence represents Rachel’s emotions on a given day. Participants were asked to estimate Rachel’s mean emotion during each day. Participants completed five practice trials, then completed the mindset induction task, and then one block of the averaging task (68 trials). Participants then completed the second mindset induction task (order was counterbalanced between participants) and performed the second and last block of the averaging task. Overall, participants completed 136 critical trials.

Results

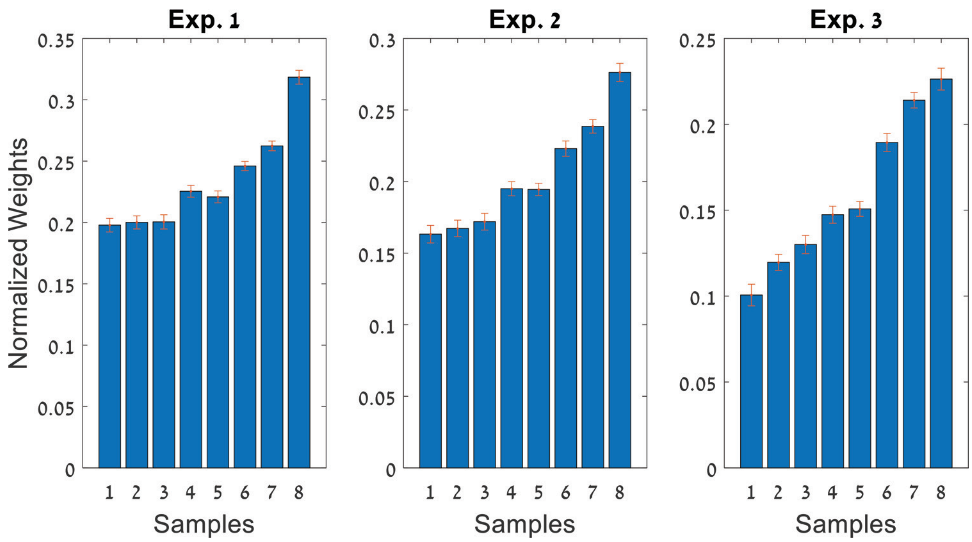
The same analysis as in Experiments 1 and 2 revealed the predicted effect of mindset, wherein Fisher-transformed correlations were higher in the abstract mindset condition (*M* = 1.008, *SD* = .196) compared to the concrete mindset condition (*M* = .955, *SD* = .191), $F(1, 58) = 5.514$, $p = .022$, $\eta_p^2 = .087$, (Figure 2C). Neither order $F(1, 58) = .607$, $p = .439$, nor the interaction of mindset and order were significant, $F(1, 58) = .914$, $p = .343$. The same pattern of results was obtained if we analyze the data using a paired-sampled permutation test ($p = .035$).

Computational Modeling

We begin by examining whether the participants’ estimations were more influenced by earlier items (i.e., primacy bias) or by later items (i.e., recency bias). Previous studies have reported the latter both for numerical values averaging task (Brezis et al., 2015; Spitzer et al., 2017), and emotional faces (Hubert-Wallander & Boynton, 2015). However, because different mechanisms may underlie these two tasks (Haberman et al., 2015; Hubert-Wallander & Boynton, 2015; Whitney & Yamanashi Leib, 2018), we analyzed the temporal weighting in each experiment separately. For each experiment, we performed a temporal regression analysis, in

Figure 4

The Temporal-Weighting Profile (Standardized Regression Weights) of the Samples in Experiments 1–3



Note. In all experiments the evaluations of the participants were more influenced by the recent items in the sequence (i.e., recency bias). Error bars correspond to within-subjects *SE*. See the online article for the color version of this figure.

which we predicted the response in each trial for Experiments 1–3 (Y), using the samples (X_i) ranked by their temporal order.

$$Y = w_o + \sum_{i=1}^{n} w_i X_i \tag{1}$$

As shown in Figure 4, across Experiments 1–3 the mean normalized weights of Samples 5–8 was higher than that of Samples 1–4 (Experiment 1: $t(54) = 7.49$, $p < .001$; Experiment 2: $t(59) = 6.97$, $p < .001$; Experiment 3: $t(59) = 9.11$, $p < .001$), indicating a recency bias. This motivated us to include a leak term (a parameter that controls the extent to which earlier values are given less weight; Usher & McClelland, 2001; Vanunu et al., 2019) in some of our models.

We examined several models to gain better understanding of the effects in Experiments 1–3. All models assume that the mean is estimated based on noisy representations of the magnitudes of the numbers (Experiments 1 and 2) or facial expressions (Experiment 3) presented in each sequence. Models are based on Equation 1 (with a Gaussian noise term), and approximate the neural population-averaging model suggested by Brezis et al. (2015, 2018; see also Rosenbaum et al., 2021 for simulations showing the equivalence between models based on Equation 1 and population-averaging models). Because similar patterns of results were obtained in Experiments 1–3, we collapsed them together.

In particular, we examined four models based on the mechanism presented in Equation 1 (see Table 1). The first model serves as a baseline, and includes equal temporal weights (i.e., no-leak) and a direct (unbiased) map of estimated values to the response scale. The second model also assumes equal weights, but includes two free-parameters: intercept and slope parameter (β_0 and β_1) which map between the participant's internal estimation to the external response scale; for slope parameters that are smaller than 1, this reflects a regression to the mean. Models 3–4 are similar to Models 1–2, but introduce leak (unequal temporal weights). For each participant, the models were fitted to the data of the abstract and concrete mindset conditions separately, and were compared with each other using the Akaike information criterion (AIC; Akaike, 1974).

As shown in Table 1, Model 4 decisively outperformed all other models both in the concrete and abstract mindset conditions, suggesting that the integration process was characterized by decreasing weight to earlier-presented samples (recency bias), as well as by biases in the mapping between presented numbers and the response scale. To examine whether these processes varied between conditions, we compared between conditions the best-fitted leak, intercept and slope parameters of Model 4.

Toward this end, we performed three *t*-tests with mindset (abstract vs. concrete) as independent variable, and leak, intercept, and slope as dependent variables.[4] The first *t* test revealed a significant difference, $t(174) = 3.48$, $p < .001$, between the slope of the abstract mindset condition ($\beta_1 = .98$) and the concrete mindset condition ($\beta_1 = .93$), indicating a more accurate, less regressive estimation mapping in the abstract condition. Similar results were obtained for the intercept parameter, $t(174) = -3.39$, $p < .001$, revealing that the intercept in the abstract mindset condition ($\beta_0 = .54$) was less biased than the intercept in the concrete mindset condition ($\beta_0 = 2.69$). Finally, while the leak was numerically lower in the concrete compared with the abstract condition, this difference was not statistically significant, $t(174) = -1.59$,

[4] For simplicity, the data in the analyses presented here were collapsed across Experiments 1–3, see SI for a full two-way ANOVAs using mindset and experiment as independent factors.

Table 1
Model Comparison

Model number	Model	Abstract mindset (aggregated AIC)	Concrete mindset (aggregated AIC)
1	$y_i = \frac{\sum_{i=1}^{n} X_i}{n} + \varepsilon_i$	97,902	98,942
2	$y_i = \beta_0 + \beta_1 \frac{\sum_{i=1}^{n} X_i}{n} + \varepsilon_i$	94,458	94,920
3	$y_i = \frac{\sum_{i=1}^{n} (1-\lambda)^{n-i} \cdot X_i}{\sum_{i=1}^{n} (1-\lambda)^{n-i}} + \varepsilon_i$	97,527	98,545
4	$y_i = \beta_0 + \beta_1 \frac{\sum_{i=1}^{n} (1-\lambda)^{n-i} \cdot X_i}{\sum_{i=1}^{n} (1-\lambda)^{n-i}} + \varepsilon_i$	**93,927**	**94,394**

Note. Note that AIC differences higher than 10 are considered decisive evidence in favor of the model with the lower numerical value (bold values indicate the best fits).

p = .12. For each condition we also computed a measure of signal-to-noise ratio, by dividing the slope by the squared root of the error-term $\left(\sqrt{\frac{\sum_{i=1}^{n} (y_{predicted} - y_{observed})^2}{n-1}}\right)$, which provides an estimation of the scatter around the regression line (Figure 2A–B). We find that this measure of signal-to-noise ratio was higher in the abstract mindset condition than in the concrete mindset condition, $t(174) = 3.83$, $p < .001$.

Experiment 4

In three experiments we have shown that abstract mindset (compared with concrete mindset) improves the accuracy by which information is aggregated in average estimations. A remaining question, however, is whether this improvement is unique to averaging, or is rather a mere reflection of a general performance improvement. Experiment 4 is a preregistered[5] examination of this question. We modified the task from average-estimation to memory for individual items. We hypothesized that abstract mindset would not improve performance in this task, and might even impair performance, compared with concrete mindset (which, compared with the abstract mindset, may confer an advantage to the processing of individual elements).

Method

Participants

One-hundred and 57 participants were recruited from Prolific[6] (79 women; M_{age} = 23.60, SD = 5.31), in exchange for $4. Participants were prescreened to those who currently hold a student status (in order to keep the sample relatively comparable to the previous lab-based experiments). Six participants were excluded, five due to below-chance performance (i.e., accuracy rate $<$ .5), and one for not completing the manipulation forms.

Materials and Procedure

Procedure generally followed that of Experiment 1 with few changes. First, instead of asking about the mean value of the sequence participants were presented after each sequence of numbers with two probe-numbers, and were asked to choose which one of them has been presented in the sequence. One of these numbers was randomly sampled from the eight numbers that were presented in that sequence, and the other number was randomly sampled from the same Gaussian distribution, but we made sure it would not be any of the numbers in the sequence. Participants completed 10 practice trials, and 80 critical trials in each mindset condition (i.e., a total of 160 critical trials), see Figure 5 for an illustration.

Second, following each critical block, participants reported their current motivation (“I was motivated to succeed in the numerical task”), stress (“I feel stressed at the moment”), and mood (“I have a positive mood at the moment”), on a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

Results

Manipulation Check

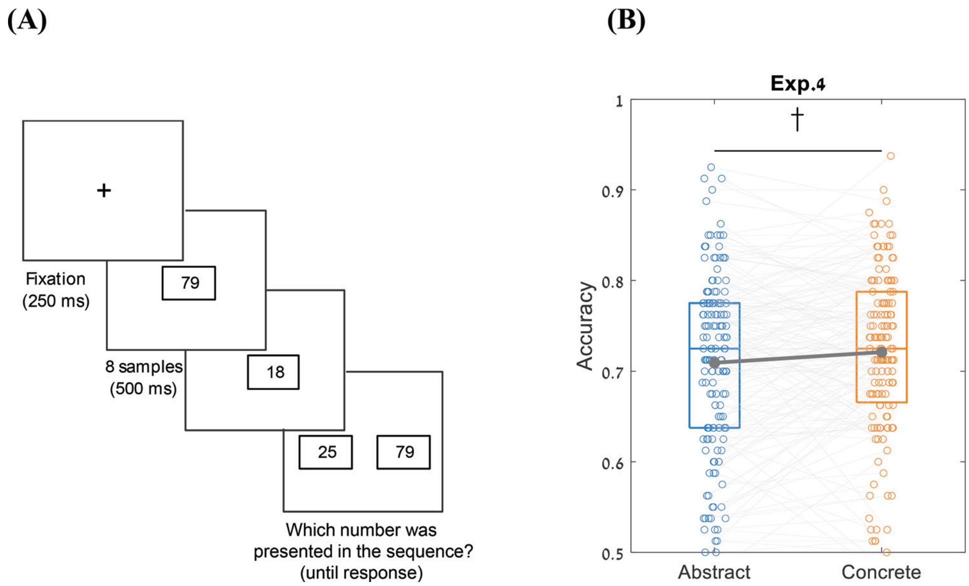
In order to validate that the manipulation elicited the intended mindset in each condition, we used an automated text analysis tool (TATE; Simchon, 2019), which rates the concreteness level of words according to a normed dictionary (Brysbaert et al., 2014). Higher scores indicate higher concreteness level. Concreteness scores were submitted to a repeated measures ANOVA, and revealed that as expected, texts’ concreteness level in the concrete mindset condition was higher (M = 2.943, SD = .321) than in the abstract mindset condition (M = 2.483, SD = .312), $F(1, 149) = 217.268$, $p < .001$, η_p^2 = .593. This analysis confirms that the experimental manipulation made participants use abstract versus concrete language, in the abstract versus concrete conditions as intended.

Accuracy

Memory-accuracy was submitted to the same analyses as in Experiments 1–3. In contrast to our previous experiments (and in contradiction to a general improvement account of our previous results), there was no advantage for the abstract condition relative to the concrete condition. Rather, the analysis revealed a marginal effect of mindset favoring the concrete mindset condition.

[5] https://aspredicted.org/blind2.php
[6] https://prolific.co/

Figure 5
Task and Results of Experiment 4



Note. (A) After each sequence of numbers, participants were presented with two probe-numbers and asked to choose which one of them has been presented in the sequence. (B) Memory accuracy rates by mindset condition. Blue and orange circles correspond to the accuracy rates of each participant in the abstract and concrete conditions, respectively. Gray circles correspond to the conditions means. The central mark on each box-plot indicates the median, and the bottom and top edges indicate the 25th and 75th percentiles, respectively. †$p < .10$. See the online article for the color version of this figure.

Accuracy in the abstract mindset condition was slightly lower (M = .709, SD = .100) than in the concrete mindset condition (M = .721, SD = .093), $F(1, 149) = 3.78$, $p = .054$, $\eta_p^2 = .025$, see Figure 5 The effect of order was not significant, $F(1, 149) = .414$, $p = .521$, but the interaction of mindset and order was significant, $F(1, 149) = 9.194$, $p = .003$, $\eta_p^2 = .058$. In addition, motivation ($p = .639$), stress ($p = .425$), and mood ($p = .189$) did not significantly differ between the two mindset conditions.

Taken together, these results indicate that the improved performance in the abstract mindset condition in Experiments 1–3 is unlikely to reflect a general effect on any type of performance (e.g., due to higher motivation, better attention, or lesser fatigue). Rather, as we discuss in the General Discussion, this pattern of findings supports the notion that abstract versus concrete processing differentially favor gist extraction versus attention to details.

General Discussion

In three experiments, we examined how abstract versus concrete thinking affected aggregation of information, and in the fourth experiment we examined how it affects memory for specific items. As predicted, we found that abstract thinking improved aggregation accuracy compared with concrete thinking, whereas it did not improve accuracy in the memory task. In Experiments 1 and 2, participants performed a numerical averaging task adopted from Brezis et al. (2015), whereas in Experiment 3 they performed a variation of the averaging task with emotional faces (Haberman et al., 2009). The results were tested using Pearson correlations as a measure of estimation precision, but similar results obtained with deviations between correct and estimated values (RMSD).[7] Finally, in Experiment 4 we ruled out a general (e.g., motivational or attentional) account for the improved performance in Experiments 1–3. When the task was changed to remembering individual items, the abstract mindset condition showed no advantage but rather a small decrement relative to the concrete mindset condition.

To better understand the processes that drive the improved average estimations in the abstract mindset, we examined several computational models, assuming that the estimation of the mean is based on transformation of numerical values (Dehaene, 1992; Dehaene & Cohen, 1991), or emotional faces (Holmes & Lourenco, 2011) into analog/magnitude representations. We found an overall recency effect—this did not differ significantly between the abstract and concrete conditions. Rather, our best fitting model indicated that concrete thinking resulted in higher levels of regression toward the mean (lower slopes) compared with abstract thinking. One way to understand the increased regression to the mean in the concrete condition is viewing it as a compensatory strategy in face of uncertainty or difficulty. This possible explanation assumes that participants in the concrete condition experienced more difficulty and/or uncertainty in their estimations and as a result provided a response that was closer to the mean of the entire block (Anobile et al., 2012;

[7] Across the three experiments, the RMSD for the abstract thinking was smaller than that of the concrete thinking, $t(174) = 2.86$, $p = .004$. In addition, there was a high correlation between the two precision measures, Pearson correlation and RMSD, across participants and conditions ($r = -.61$, $p < .001$)

Jazayeri & Shadlen, 2010). Indeed, previous work suggests that participants have access to their averaging uncertainty as their estimation confidence decreases with the variance of the sequence (Rosenbaum et al., 2021). Future studies could examine the specific strategy that gives rise to more regressive responses in the concrete processing condition: Is it the case that participants in this condition represent (a limited number of) individual samples rather than keeping track of a "running average"? Do they fail to keep track of the boundaries between different trials? Do they provide a "default" regressive middle-of-the-scale response on a larger proportion of the trials? All of these (nonmutually exclusive) strategies would produce more regressive responses in the concrete than the abstract condition.

Finally, the measure of signal-to-noise ratio (slope/error-term) was higher for the abstract condition, indicating that abstract thinking created a less noisy representation of the sequences of values in each trial. We find this pattern of results noteworthy, because it indicates that abstract processing gave rise to better, more accurate integration, that at the same time was not characterized by increased information loss (as would be predicted, e.g., if abstraction would lead to heuristic, less elaborate processing).

Another prediction derived from CLT is that an abstract mindset would improve performance when gist extraction or filtering-out irrelevant information is beneficial (e.g., Hadar et al., 2019), but not when one needs to retain the exact details. The results of Experiment 4 are consistent with this prediction. However, future research with a larger sample of participants and stimuli types will be needed to examine this hypothesis more closely.

In conclusion, people may face a need to average series of values when they consider which stock to buy, which people to befriend, or which student shows more promising performance. Understanding the contextual features that contribute to the accuracy of the averaging process could thus inform not only theories of decision-making and (social) cognition, but also educators and policymakers.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Anderson, N. H. (1971). Integration theory and attitude change. *Psychological Review*, *78*(3), 171–206. https://doi.org/10.1037/h0030834

Anderson, N. H. (1981). *Foundations of information integration theory* (Vol. 578). Academic Press.

Anobile, G., Cicchini, G. M., & Burr, D. C. (2012). Linear mapping of numbers onto space requires attention. *Cognition*, *122*(3), 454–459. https://doi.org/10.1016/j.cognition.2011.11.006

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*(2), 157–162. https://doi.org/10.1111/1467-9280.00327

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, *41*, 258–290. https://doi.org/10.1037/h0055756

Baek, J., & Chong, S. C. (2020). Ensemble perception and focused attention: Two different modes of visual processing to cope with limited capacity. *Psychonomic Bulletin & Review*, *27*(4), 602–606. https://doi.org/10.3758/s13423-020-01718-7

Betsch, T., Plessner, H., Schwieren, C., & Gütig, R. (2001). I like it but I don't know why: A value-account approach to implicit attitude formation. *Personality and Social Psychology Bulletin*, *27*(2), 242–253. https://doi.org/10.1177/0146167201272009

Brezis, N., Bronfman, Z. Z., & Usher, M. (2015). Adaptive spontaneous transitions between two mechanisms of numerical averaging. *Scientific Reports*, *5*(1), 10415. https://doi.org/10.1038/srep10415

Brezis, N., Bronfman, Z. Z., Jacoby, N., Lavidor, M., & Usher, M. (2016). Transcranial direct current stimulation over the parietal cortex improves approximate numerical averaging. *Journal of Cognitive Neuroscience*, *28*(11), 1700–1713. https://doi.org/10.1162/jocn_a_00991

Brezis, N., Bronfman, Z. Z., & Usher, M. (2018). A perceptual-like population-coding mechanism of approximate numerical averaging. *Neural Computation*, *30*(2), 428–446. https://doi.org/10.1162/neco_a_01037

Brusovansky, M., Glickman, M., & Usher, M. (2018). Fast and effective: Intuitive processes in complex decisions. *Psychonomic Bulletin & Review*, *25*(4), 1542–1548. https://doi.org/10.3758/s13423-018-1474-1

Brusovansky, M., Vanunu, Y., & Usher, M. (2019). Why we should quit while we're ahead: When do averages matter more than sums? *Decision*, *6*(1), 1–16. https://doi.org/10.1037/dec0000087

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Burgoon, E. M., Henderson, M. D., & Markman, A. B. (2013). There are many ways to see the forest for the trees: A tour guide for abstraction. *Perspectives on Psychological Science*, *8*(5), 501–520. https://doi.org/10.1177/1745691613497964

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*(4), 393–404. https://doi.org/10.1016/S0042-6989(02)00596-5

Chong, S. C., & Treisman, A. (2005). Statistical processing: Computing the average size in perceptual groups. *Vision Research*, *45*(7), 891–900. https://doi.org/10.1016/j.visres.2004.10.004

Chong, S. C., Joo, S. J., Emmanouil, T. A., & Treisman, A. (2008). Statistical processing: Not so implausible after all. *Perception & Psychophysics*, *70*(7), 1327–1334. https://doi.org/10.3758/pp.70.7.1327

Dakin, S. C., & Watt, R. J. (1997). The computation of orientation statistics from visual texture. *Vision Research*, *37*(22), 3181–3192. https://doi.org/10.1016/S0042-6989(97)00133-8

de Fockert, J., & Wolfenstein, C. (2009). Rapid extraction of mean identity from sets of faces. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *62*(9), 1716–1722. https://doi.org/10.1080/17470210902811249

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*(1-2), 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S., & Cohen, L. (1991). Two mental calculation systems: A case study of severe acalculia with preserved approximation. *Neuropsychologia*, *29*(11), 1045–1054. https://doi.org/10.1016/0028-3932(91)90076-K

Ding, Y., & Keh, H. T. (2017). Consumer reliance on intangible versus tangible attributes in service evaluation: The role of construal level. *Journal of the Academy of Marketing Science*, *45*(6), 848–865. https://doi.org/10.1007/s11747-017-0527-8

Efrat-Treister, D., Daniels, M. A., & Robinson, S. L. (2020). Putting time in perspective: How and why construal level buffers the relationship between wait time and aggressive tendencies. *Journal of Organizational Behavior*, *41*(3), 294–309. https://doi.org/10.1002/job.2433

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior*, *1*(1), 56–75. https://doi.org/10.1007/BF01115465

Eyal, T., Hoover, G. M., Fujita, K., & Nussbaum, S. (2011). The effect of distance-dependent construals on schema-driven impression formation. *Journal of Experimental Social Psychology*, *47*(1), 278–281. https://doi.org/10.1016/j.jesp.2010.10.007

Florey, J., Clifford, C. W., Dakin, S., & Mareschal, I. (2016). Spatial limitations in averaging social cues. *Scientific Reports*, *6*(1), 32210. https://doi.org/10.1038/srep32210

Freitas, A. L., Gollwitzer, P., & Trope, Y. (2004). The influence of abstract and concrete mindsets on anticipating and guiding others' self-regulatory efforts. *Journal of Experimental Social Psychology*, *40*(6), 739–752. https://doi.org/10.1016/j.jesp.2004.04.003

Gilead, M., Liberman, N., & Maril, A. (2014). From mind to matter: Neural correlates of abstract and concrete mindsets. *Social Cognitive and Affective Neuroscience*, *9*(5), 638–645. https://doi.org/10.1093/scan/nst031

Gilead, M., Trope, Y., & Liberman, N. (2020). Above and beyond the concrete: The diverse representational substrates of the predictive brain. *Behavioral and Brain Sciences*, *43*, e121. https://doi.org/10.1017/S0140525X19002000

Haberman, J., Brady, T. F., & Alvarez, G. A. (2015). Individual differences in ensemble perception reveal multiple, independent levels of ensemble representation. *Journal of Experimental Psychology: General*, *144*(2), 432–446. https://doi.org/10.1037/xge0000053

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*, *17*(17), R751–R753. https://doi.org/10.1016/j.cub.2007.06.039

Haberman, J., Harp, T., & Whitney, D. (2009). Averaging facial expression over time. *Journal of Vision*, *9*(11), 1. https://doi.org/10.1167/9.11.1

Hadar, B., Luria, R., & Liberman, N. (2019). Concrete mindset impairs filtering in visual working memory. *Psychonomic Bulletin & Review*, *26*(6), 1917–1924. https://doi.org/10.3758/s13423-019-01625-6

Hadar, B. (2021, June 2). *Abstract thinking facilitates aggregation of information*. osf.io/y4uxd/

Hamilton, D. L., & Sherman, S. J. (1996). Perceiving persons and groups. *Psychological Review*, *103*(2), 336–355. https://doi.org/10.1037/0033-295X.103.2.336

Hansen, J., & Trope, Y. (2013). When time flies: How abstract and concrete mental construal affect the perception of time. *Journal of Experimental Psychology: General*, *142*(2), 336–347. https://doi.org/10.1037/a0029283

Holmes, K. J., & Lourenco, S. F. (2011). Common spatial organization of number and emotional expression: A mental magnitude line. *Brain and Cognition*, *77*(2), 315–323. https://doi.org/10.1016/j.bandc.2011.07.002

Hubert-Wallander, B., & Boynton, G. M. (2015). Not all summary statistics are made equal: Evidence from extracting summaries across time. *Journal of Vision*, *15*(4), 5. https://doi.org/10.1167/15.4.5

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

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus & Giroux.

Khayat, N., & Hochstein, S. (2018). Perceiving set mean and range: Automaticity and precision. *Journal of Vision*, *18*(9), 23. https://doi.org/10.1167/18.9.23

Kille, D. R., Eibach, R. P., Wood, J. V., & Holmes, J. G. (2017). Who can't take a compliment? The role of construal level and self-esteem in accepting positive feedback from close others. *Journal of Experimental Social Psychology*, *68*(1), 40–49. https://doi.org/10.1016/j.jesp.2016.05.003

Liberman, N., & Trope, Y. (2014). Traversing psychological distance. *Trends in Cognitive Sciences*, *18*(7), 364–369. https://doi.org/10.1016/j.tics.2014.03.001

Malmi, R. A., & Samson, D. J. (1983). Intuitive averaging of categorized numerical stimuli. *Journal of Verbal Learning and Verbal Behavior*, *22*(5), 547–559. https://doi.org/10.1016/S0022-5371(83)90337-7

Maule, J., Witzel, C., & Franklin, A. (2014). Getting the gist of multiple hues: Metric and categorical effects on ensemble perception of hue. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *31*(4), A93–A102. https://doi.org/10.1364/JOSAA.31.000A93

Napier, J. L., Luguri, J. B., Dovidio, J. F., & Oltman, K. A. (2018). Construing the essence: The effects of construal level on genetic attributions for individual and social group differences. *Personality and Social Psychology Bulletin*, *44*(9), 1395–1407. https://doi.org/10.1177/0146167218768799

Neumann, M. F., Schweinberger, S. R., & Burton, A. M. (2013). Viewers extract mean and individual identity from sets of famous faces. *Cognition*, *128*(1), 56–63. https://doi.org/10.1016/j.cognition.2013.03.006

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, *4*(7), 739–744. https://doi.org/10.1038/89532

Post, R. B., Haberman, J., Iwaki, L., & Whitney, D. (2012). The frozen face effect: Why static photographs may not do you justice. *Frontiers in Psychology*, *3*(1), 22. https://doi.org/10.3389/fpsyg.2012.00022

Robitaille, N., & Harris, I. M. (2011). When more is less: Extraction of summary statistics benefits from larger sets. *Journal of Vision*, *11*(12), 18. https://doi.org/10.1167/11.12.18

Rosenbaum, D., Glickman, M., & Usher, M. (2021). Extracting summary statistics of rapid numerical sequences. *Frontiers in Psychology*, *12*, 693575. https://doi.org/10.3389/fpsyg.2021.693575

Simchon, A. (2019). TATE: Text analysis tools for English (R package version 3.0) [Computer software]. https://github.com/almogsi/TATE

Solomon, J. A. (2010). Visual discrimination of orientation statistics in crowded and uncrowded arrays. *Journal of Vision*, *10*(14), 19. https://doi.org/10.1167/10.14.19

Spitzer, B., Waschke, L., & Summerfield, C. (2017). Selective overweighting of larger magnitudes during noisy numerical comparison. *Nature Human Behaviour*, *1*(8), 145. https://doi.org/10.1038/s41562-017-0145

Spunt, R. P., Kemmerer, D., & Adolphs, R. (2016). The neural basis of conceptualizing the same action at different levels of abstraction. *Social Cognitive and Affective Neuroscience*, *11*(7), 1141–1151. https://doi.org/10.1093/scan/nsv084

Stillman, P. E., Medvedev, D., & Ferguson, M. J. (2017). Resisting temptation: Tracking how self-control conflicts are successfully resolved in real time. *Psychological Science*, *28*(9), 1240–1258. https://doi.org/10.1177/0956797617705386

Sweeny, T. D., & Whitney, D. (2014). Perceiving crowd attention: Ensemble perception of a crowd's gaze. *Psychological Science*, *25*(10), 1903–1913. https://doi.org/10.1177/0956797614544510

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440–463. https://doi.org/10.1037/a0018963

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550–592. https://doi.org/10.1037/0033-295X.108.3.550

van Osch, Y., Blanken, I., Meijs, M. H., & van Wolferen, J. (2015). A group's physical attractiveness is greater than the average attractiveness of its members: The group attractiveness effect. *Personality and Social Psychology Bulletin*, *41*(4), 559–574. https://doi.org/10.1177/0146167215572799

Vanunu, Y., Pachur, T., & Usher, M. (2019). Constructing preference from sequential samples: The impact of evaluation format on risk attitudes. *Decision*, *6*(3), 223–236. https://doi.org/10.1037/dec0000098

Walker, D., & Vul, E. (2014). Hierarchical encoding makes individuals in a group seem more attractive. *Psychological Science*, *25*(1), 230–235. https://doi.org/10.1177/0956797613497969

Watamaniuk, S. N., & Duchon, A. (1992). The human visual system averages speed information. *Vision Research*, *32*(5), 931–941. https://doi.org/10.1016/0042-6989(92)90036-I

Watamaniuk, S. N., & McKee, S. P. (1998). Simultaneous encoding of direction at a local and global scale. *Perception & Psychophysics*, *60*(2), 191–200. https://doi.org/10.3758/BF03206028

Watamaniuk, S. N., Sekuler, R., & Williams, D. W. (1989). Direction perception in complex dynamic displays: The integration of direction information. *Vision Research*, *29*(1), 47–59. https://doi.org/10.1016/0042-6989(89)90173-9

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology*, *69*(1), 105–129. https://doi.org/10.1146/annurev-psych-010416-044232

Yamanashi Leib, A., Chang, K., Xia, Y., Peng, A., & Whitney, D. (2020). Fleeting impressions of economic value via summary statistical representations. *Journal of Experimental Psychology: General*, *149*(10), 1811–1822. https://doi.org/10.1037/xge0000745

Yamanashi Leib, A., Kosovicheva, A., & Whitney, D. (2016). Fast ensemble representations for abstract visual impressions. *Nature Communications*, *7*(1), 13186. https://doi.org/10.1038/ncomms13186

Yudkin, D. A., Liberman, N., Wakslak, C., & Trope, Y. (2020). Better off and far away: Reactions to others' outcomes depends on their distance. *Organizational Behavior and Human Decision Processes*, *156*, 13–23. https://doi.org/10.1016/j.obhdp.2019.10.004

Received September 7, 2020
Revision received June 1, 2021
Accepted July 9, 2021 ■