The scarcity heuristic impacts reward processing within the medial-frontal cortex

Chad C. Williams[a], Boaz Y. Saffer[b], Robert B. McCulloch[a] and Olave E. Krigolson[a]

Objects that are rare are often perceived to be inherently more valuable than objects that are abundant – a bias brought about in part by the scarcity heuristic. In the present study, we sought to test whether perception of rarity impacted reward evaluation within the human medial-frontal cortex. Here, participants played a gambling game in which they flipped rare and abundant 'cards' on a computer screen to win financial rewards while electroencephalographic data were recorded. Unbeknownst to participants, reward outcome and frequency was random and equivalent for both rare and abundant cards; thus, only a perception of scarcity was true. Analysis of the electroencephalographic data indicated that the P300 component of the event-related brain potential differed in amplitude for wins and losses following the selection of rare cards, but not following the selection of abundant cards. Importantly, then, we found that the perception of card rarity impacted reward processing even though reward feedback was independent of and subsequent to card selection. Our data indicate a top-down influence of the scarcity heuristic on reward evaluation, and specifically the processing of reward magnitude, within the human medial-frontal cortex. *NeuroReport* **27:522–526** **Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.**

NeuroReport 2016, **27**:522–526

Keywords: electroencephalography, feedback, medial-frontal cortex, heuristic, P300, reinforcement learning, reward, reward positivity, scarcity, value

[a]Neuroeconomics Laboratory, Department of Exercise Science, Physical & Health Education, University of Victoria, Victoria and [b]Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada

Correspondence to Chad C. Williams, School of Exercise Science, Physical & Health Education, University of Victoria, PO. Box 17000 STN CSC, Victoria, BC, Canada V8W 2Y2
Tel: + 1 250 721 8373; fax: + 1 250 721 6601; e-mail: ccwillia@uvic.ca

Received 16 February 2016 **accepted** 2 March 2016

Introduction

To make decisions, we frequently rely on heuristics, or rules of thumb, to solve complex problems [1]. For instance, when attempting to determine the value of unknown items, we may utilize the scarcity heuristic – an assumption that rare items are inherently more valuable than abundant items simply because they are rare. A classic example of the scarcity heuristic is the valuation of rare marbles: children afford white marbles more value simply because white marbles are quite rare, even though they are equivalent in financial value to other marbles. The marble example is backed up by a multitude of behavioural studies that show that scarce items are perceived as being more valuable than abundant items, even though there is no evidence to justify this supposition [2–5]. For example, in one study, Parker and Lehmann [4] found that when participants were presented with an array of two different wines to pick from, they more often chose a wine bottle that was scarce rather than one that was abundant, even though there was no other basis for their decision. Although there is a large body of behavioural evidence documenting the scarcity heuristic, little is known about how the scarcity heuristic biases decision-making processes within the brain.

At the neural level, decision-making processes are dependent on learning systems within the brain to learn values to optimize response selection [6–9]. Neural learning systems are typically thought to have at least two key (and perhaps more) components that are temporally independent [10,11]. For example, studies using electroencephalography (EEG) have found that the reward positivity – a component of the human event-related brain potential (ERP) associated with reward evaluation – reflects an early reward evaluation process sensitive to the discrepancy between outcomes and expectations (i.e. 'prediction errors' [6–9]). Subsequent to the reward positivity, late reward processes are more focused on the allocation of attentional resources to facilitate the updating of reward contingencies and are thought to be indexed by the P300 ERP component. The P300 typically occurs 400–600 ms following a visual stimulus and has been shown to be sensitive to the updating of internal models and/or reward valence [10–13]. Given that the scarcity heuristic is believed to bias perception of value, one may posit that the scarcity heuristic biases early, late or both levels of reward processing.

In the present study, we sought to investigate the effects of the scarcity heuristic on reward evaluation within the medial-frontal cortex. Participants completed a computerized gambling task, wherein they 'flipped' scarce and abundant cards to win financial rewards while EEG data were recorded. Unbeknownst to participants, their wins and losses were independent of card selection – gambling outcomes were random. We hypothesized that

0959-4965 Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved. DOI: 10.1097/WNR.0000000000000575

Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.

participants would be biased by the perception of card scarcity and that this would in turn bias reward evaluation. Specifically, we predicted that gambling outcomes following the selection of rare cards would be biased by the scarcity heuristic – a result that we would observe as a differential modulation of one or more components of the human ERP known to be associated with reward processing – the reward positivity [6–8] and/or the P300 [10–14]. However, our initial belief was that the perception of card scarcity would not bias early reward processing. Indeed, we hypothesized that the amplitude of the reward positivity would not show a scarcity bias, given that previous work has shown a lack of sensitivity of this component to reward value [6,15,16].

Materials and methods

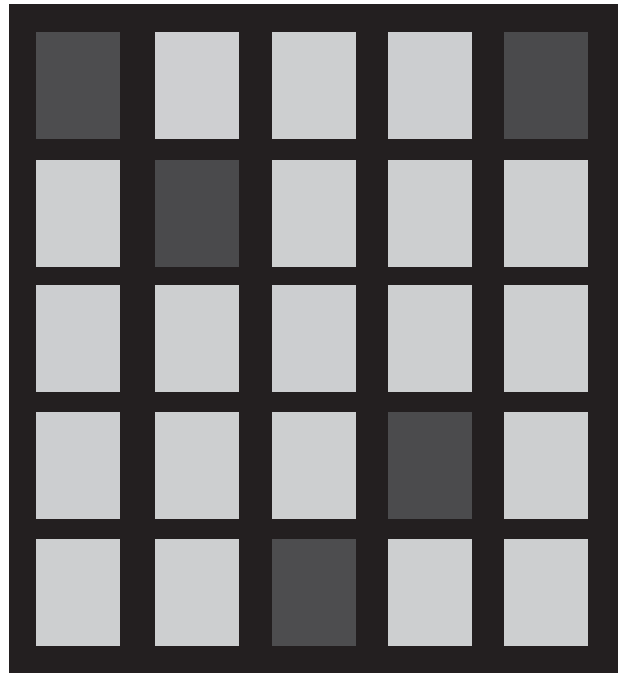
Participants

Seventeen right-handed undergraduate students (10 women, age range 18–25 years) with no known neurological impairments and normal or corrected-to-normal vision were recruited using the UBC Psychology Research Participation System. However, we excluded three of the 17 participants completely from further analysis because they did not follow task instructions. Participants provided informed consent as approved by the UBC Office of Research Ethics: Human Ethical Review, and the study was carried out in accordance with the ethical standards as prescribed in the 1964 Declaration of Helsinki. Participants were compensated by obtaining three course credits and a monetary bonus on the basis of a proportion of their total winnings (~ \$10 on average).

Apparatus and procedure

Participants were seated in a sound-dampened room in front of a 19″ LCD computer monitor and used a standard USB mouse to perform a gambling task [written in Matlab (version 8.4; Mathworks, Natick, Massachusetts, USA)] using the Psychophysics Toolbox extension [17]. At the start of each trial of the gambling task, participants viewed a five by five array of coloured 'cards' (squares) on a black background (Fig. 1). The cards were either one of two randomly generated colours, with one colour being scarce and the other being abundant. The number of scarce cards (n: ~20%) was determined from a Gaussian distribution ($M=5$, $SD=1$); the rest of the cards ($25-n$: ~ 80%) were considered abundant. Note, however, no less than three scarce cards were ever presented for methodological reasons (see below). Following initial viewing of the card array, participants were instructed to select six cards one at a time with the constraint that they had to select three cards of each colour. Following the selection of each card, a white fixation cross appeared within the centre of the selected card for 1000–1200 ms. Subsequent to the fixation cross, the card 'flipped' and a reward feedback stimulus appeared for 1000 ms. Reward feedback was either a '\$' or a '0', indicating, respectively, a win or a loss (equivalent probability of either outcome). After presentation of reward feedback, the card disappeared from the screen and participants were prompted to select another card. This process continued until six cards had been selected in total – three that were scarce and three that were abundant. Thus, our methodology created four feedback conditions: scarce win, scarce loss, abundant win and abundant loss. Information on the amount won from each gamble was not presented after each selection, but instead, at the end of the trial. The total reward payout for each trial was generated using a random number from a Gaussian distribution ($M=80$, $SD=10$) to prevent the participants from learning that reward outcomes were random and independent of card selection. Participants completed 40 trials for a total of 240 individual gambles (120 scarce, 120 abundant).

Fig. 1



Display of the task when participants had to choose which 'card' (square) to gamble on. This display included an array of 25 cards, each of which are one of two colours. One of the colours was rare (e.g. dark grey) and the other was abundant (e.g. light grey).

Data acquisition

The EEG was recorded from 40 electrode locations using ActiView software (Biosemi B.V., Amsterdam, The Netherlands). The electrodes were mounted in a fitted cap with a standard 10–20 layout and were referenced to a two-electrode feedback loop (common mode sense to driven right leg). The vertical and horizontal electro-oculograms were recorded from electrodes placed above and below the right eye and on the outer canthi of the left and right eyes, respectively. Electrode offsets were

Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.

maintained below ± 25 mV at all times. The EEG data were sampled at 256 Hz and amplified using an Active Two system (Biosemi B.V., Amsterdam, The Netherlands).

Data analysis

Offline EEG data analyses were carried out using Brain Vision Analyzer 2.0 software (Brainproducts GmbH, Munich, Germany). First, channels that were determined to have excessive artefacts and/or noise were excluded from analysis. Next, the EEG data were re-referenced to an average mastoid reference and filtered using a dual-pass Butterworth filter (0.1–30 Hz; also a 60 Hz notch filter was applied). Segments of data 3000 ms in length were then extracted from the continuous EEG locked to the onset of each occurrence of a feedback stimulus. Each segment spanned from 1000 ms before the event of interest to 2000 ms after; epoch length was chosen to facilitate independent component analysis. Independent component analysis was then used to remove ocular artefacts [18]. Following the independent component analysis, missing channels were interpolated using spherical splines. Data were then resegmented into shorter epochs spanning from 200 ms before to 600 ms after each event of interest for each of the feedback conditions: scarce win, scarce loss, abundant win and abundant loss. Subsequent to this, a baseline correction was implemented using the 200 ms before feedback stimuli onset for each segment. Next, segments for each of the four feedback conditions were subjected to an artefact rejection algorithm with a 10 μV/ms gradient and 150 μV absolute difference criteria. The artefact algorithm resulted in an average of 1.0% (0.2–1.7%) of the data being excluded from further analysis.

Average ERP waveforms were created for each participant by averaging the segmented EEG data for each electrode and each condition (scarce win, scarce loss, abundant win, abundant loss). Difference waveforms for each participant were then constructed by subtracting average loss waveforms from average win waveforms for both the scarce and the abundant conditions [18]. Finally, for each feedback condition and the two difference waveforms, we constructed grand average waveforms by averaging corresponding waveforms across all participants.

Given our hypotheses, we focused subsequent analyses on two components of interest: the reward positivity and the P300. We quantified the reward positivity for each participant and condition (scarce, abundant) as the maximal positive deflection on subject difference waveforms at channel FCz within a 100 ms window surrounding the peak difference of the reward positivity on the appropriate grand average difference waveform (Table 1). We focused our analysis on channel FCz and a 100 ms time window because of visual inspection and previous literature [6–8]. We quantified P300 amplitude using the same process, but at a different electrode site, channel Pz (Table 1). Again, our quantification was based on visual inspection and previous research [10–13].

Table 1 Average peak times with 95% confidence intervals of the reward positivity at channel FCz and the P300 at channel Pz, with intervals of analysis for both conditions

Components	Condition	Peak time (ms)	95% CI (±) (ms)	Minimum interval (ms)	Maximum interval (ms)
Reward positivity	Scarce	320	9	270	370
	Abundant	319	13	269	369
P300	Scarce	405	15	355	455
	Abundant	399	14	349	449

Intervals for the reward positivity and the P300 were ± 50 ms centred on the average peak times.
CI, confidence interval.

Single-sample t-tests were used to test for component existence [8] and paired-samples t-tests were used to compare condition effects. Component existence analyses were carried out by determining whether there was a significant difference between win and loss waveforms for each component in each condition. Difference waveform analyses were carried out by determining whether there was a significant difference between component amplitudes across conditions. An α level of 0.05 was assumed for all statistical tests. Error measures for descriptive and inferential statistics reflect 95% confidence intervals [19].

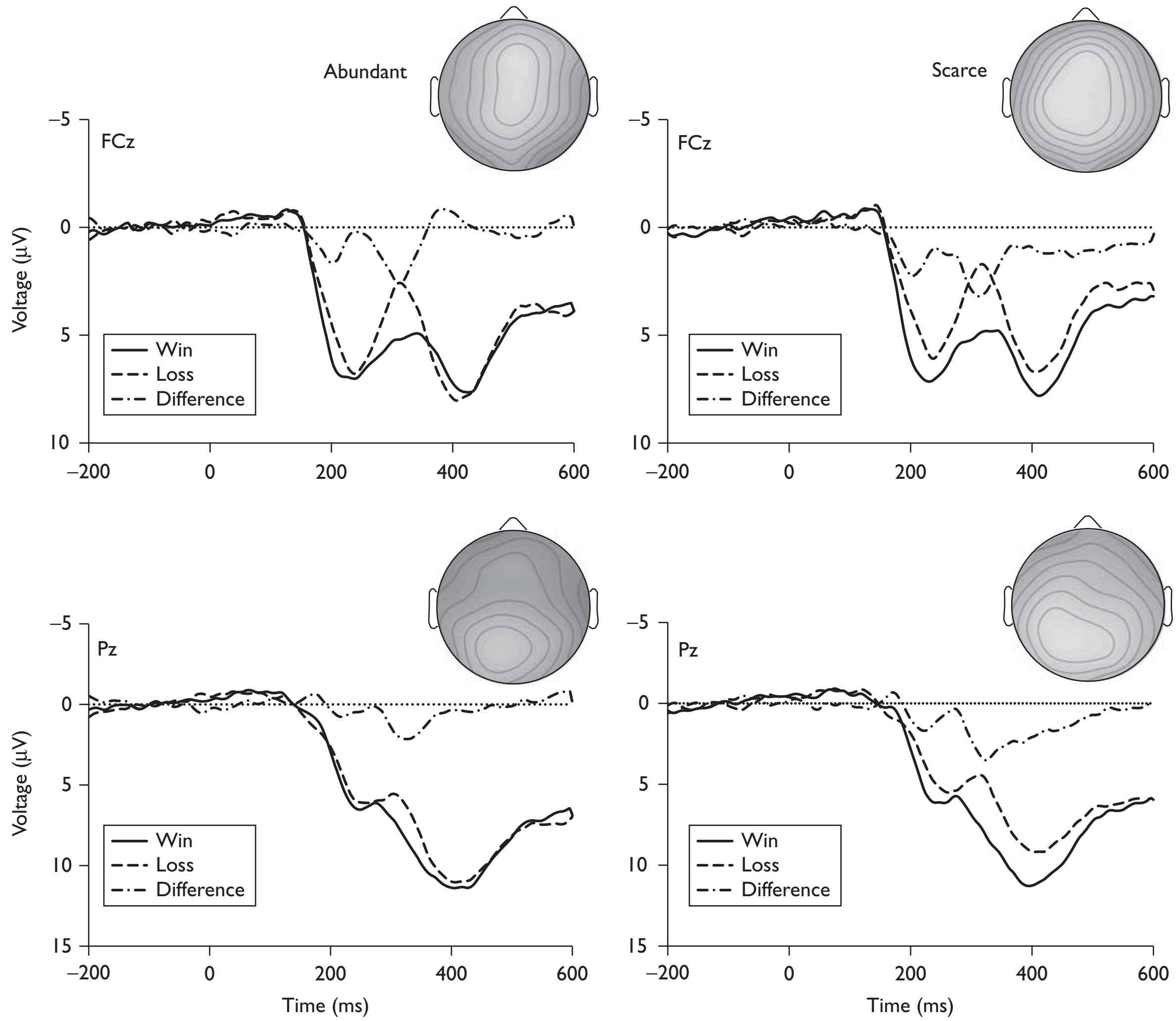
Results

Our analysis of the grand average ERP waveforms indicated components with scalp topography and timing consistent with the reward positivity and the P300 (Fig. 2). The components were quantified as the minimal (reward positivity) or the maximal (P300) measure of each deflection. The reward positivity was maximal at front-central areas of the scalp (i.e. channel FCz) and the P300 was maximal at parietal-central areas of the scalp (i.e. channel PZ) for both conditions. Single-sample t-tests showed component existence for the reward positivity [scarce: $t(13)=3.27$, $P=0.006$; abundant: $t(13)=2.37$, $P=0.034$] and a difference in the P300 between wins and losses in the scarce condition, $t(13)=2.88$, $P=0.013$. A single-sample t-test indicated that the P300 did not differ between wins and losses in the abundant condition, $t(13)=0.19$, $P=0.850$.

Experimental condition (scarce, abundant) did not impact the amplitude of the reward positivity [2.92 μV (2.00–3.85 μV) vs. 2.31 μV (0.99–3.64 μV)]. Specifically, our statistical comparison showed that the amplitude of the scarce reward positivity and the abundant reward positivity did not differ, $t(13)=0.99$, $P=0.341$, $d=0.26$, effect = 0.61 μV (− 0.72 to 1.94 μV). However, we did observe that the amplitude of P300 was impacted by

Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.

Fig. 2



Conditional waveforms, difference waveforms and topographic maps for both conditions. Top: analysis of the reward positivity; bottom: analysis of the P300; left: abundant condition; right: scarce condition. Difference waveforms were created for each component by subtracting the abundant condition from the scarce condition. Scalp distributions for the reward positivity were created by subtracting the abundant condition from the scarce condition and averaging the peak topographies of all participants. The potential scale ranges from −4.0 μV (white) to 4.0 μV (dark grey). Scalp distributions for the P300 were created by subtracting the abundant condition from the scarce condition and averaging the peak topographies of all participants. The potential scale ranges from −5.0 μV (white) to 5.0 μV (dark grey).

experimental condition [scarce: 4.81 μV (2.94–6.69 μV) vs. abundant: 3.31 μV (1.67–4.95 μV)]. Specifically, our statistical comparison showed that the P300 was larger for gambles in the scarce as opposed to the abundant condition, $t(13)=3.11$, $P=0.008$, $d=0.83$, effect = 1.50 μV (0.46–2.54 μV).

Discussion

In the present study, we have shown that the scarcity heuristic biased human reward processing. We found that the amplitude of the P300, an ERP component that has been shown to be sensitive to reward magnitude, was enhanced for gambles that were perceived to be scarce relative to gambles that were perceived to be abundant – a result congruent with studies that have shown that scarcity enhances the perceived value of objects [2–5]. It is important to note here that the scarcity bias carried over from gamble selection, wherein we did bias perception of scarcity to gamble outcome that was actually equivalent and equiprobable, thus negating potential ERP frequency confounds [15]. In line with our hypotheses, we also found that early reward processing (i.e. the reward positivity) was not impacted by perception of scarcity – a result that makes sense, given that the amplitude of the reward positivity is typically considered to be insensitive to reward magnitude [6,15,16]. Our study thus also supports a dissociation between early and late reward processing. Specifically, the differential bias

Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.

of the scarcity heuristic on late but not early reward processing supports the independence of these two processes – a result in line with previous findings [12,13].

A multitude of behavioural studies have shown that the scarcity heuristic impacts perception of value [2–5]. Here, we show that heuristics bias neural processes – our data provide evidence of a top-down influence of the scarcity heuristic on reward processing. A key problem with reinforcement learning (RL) explanations for human learning is that RL solutions are very slow in complex environments [9]. For example, RL solutions for the game of chess are impractical without modification. One possible way to modify RL solutions to 'speed them up' when faced with complex problems is through the use of heuristics. Indeed, heuristics by definition are 'short-cuts' that facilitate our ability to learn quickly and thus make more rapid decisions [1]. Importantly, our data support this contention and suggest that heuristics do influence reward processing, perhaps in part to speed up learning and/or complex decision processes.

Interestingly, here, we observed that the bias of scarcity was independent from others – we observed a scarcity bias for a participant performing a task in isolation. Specifically, we observed the scarcity bias in a study where the influence of others did not affect the participant's current selection, nor did the participant's response affect the selection of others. As the scarcity effect was still observed in our study, our data suggest that the bias is actually independent of the influence of others. As such, our results are in contrast to claims that this bias may be driven by external factors such as popularity [4] and uniqueness from others [20]. Instead, our results indicate that the scarcity bias must be derived by more internal factors, for example – increased attention [21].

An alternative explanation of our findings could be that the process used to manipulate reward magnitude (scarcity) affected RL differently than more traditional methods (e.g. presenting different magnitudes of rewards). As indicated above, our results suggest that the scarcity bias is driven by internal factors, possibly increased attention [21]. Therefore, the effect that we observed could rather be an effect of attention rather than magnitude. The P300 is indeed influenced by attention [22,23]; however, this alternative seems unlikely because studies have also found attention to modulate the reward positivity [24,25]. Therefore, we believe that our data do support a top-down influence of the scarcity heuristic on reward processing.

Conclusion

Our results support the claim that top-down processes such as the scarcity heuristic enhance the perceived value of items as we found that P300 amplitude was differentially modulated by perceived scarcity. Furthermore, we found that the scarcity bias extends to late, but not early reward processing, supporting the independence of these two components of reward evaluation.

Acknowledgements

This study was funded by Natural Sciences and Engineering Research Council of Canada (NSERC) and the Neuroeducation Network.

Conflicts of interest

There are no conflicts of interest.

References

1 Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. *Science* 1974; **185**:1124–1131.
2 Effron DA, Miller DT. Diffusion of entitlement: an inhibitory effect of scarcity on consumption. *J Exp Soc Psychol* 2011; **47**:378–383.
3 Mittone L, Savadori L. The scarcity bias. *Appl Psychol* 2009; **58**:453–468.
4 Parker JR, Lehmann DR. When shelf-based scarcity impacts consumer preferences. *J Retailing* 2011; **87**:142–155.
5 Suri R, Kohli C, Monroe KB. The effects of perceived scarcity on consumers' processing of price information. *J Aca Mark Sci* 2007; **35**:89–100.
6 Holroyd CB, Coles MGH. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 2002; **109**:679–709.
7 Holroyd CB, Krigolson OE. Reward prediction error signals associated with a modified time estimation task. *Psychophysiology* 2007; **44**:913–917.
8 Krigolson OE, Hassall CD, Handy TC. How we learn to make decisions: rapid propagation of reinforcement learning prediction errors in humans. *J Cogn Neurosci* 2014; **26**:635–644.
9 Sutton RS, Barto AG. *Reinforcement learning: an introduction*. Cambridge, MA: MIT Press; 1998.
10 Luu P, Tucker DM, Stripling R. Neural mechanisms for learning actions in context. *Brain Res* 2007; **1179**:89–105.
11 Luu P, Shane M, Pratt NL, Tucker DM. Corticolimbic mechanisms in the control of trial and error learning. *Brain Res* 2009; **1247**:100–113.
12 Sato A, Yasuda A, Ohira H, Miyawaki K, Nishikawa M, Kumano H, *et al.* Effects of value and reward magnitude on feedback negativity and P300. *Neuroreport* 2005; **16**:407–411.
13 Yeung N, Sanfey AG. Independent coding of reward magnitude and valence in the human brain. *J Neurosci* 2004; **24**:6258–6264.
14 Goyer J, Woldorff M, Huettel S. Rapid electrophysiological brain responses are influenced by both valence and magnitude of monetary rewards. *J Cogn Neurosci* 2008; **20**:2058–2069.
15 Holroyd CB, Pakzad-Vaezi KL, Krigolson OE. The feedback correct-related positivity: sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology* 2008; **45**:688–697.
16 Proudfit GH. The reward positivity: from basic research on reward to a biomarker for depression. *Psychophysiology* 2015; **52**:449–459.
17 Brainard DH. The psychophysics toolbox. *Spat Vis* 1997; **10**:433–436.
18 Luck SJ. *An introduction to the event-related potential technique*, 2nd ed. Cambridge, MA: MIT Press; 2014.
19 Cumming G. The new statistics: a how-to guide. *Aust Psychol* 2013; **48**:161–170.
20 Snyder CR. Product scarcity by need for uniqueness interaction: a consumer catch-22 carousel? *Basic Appl Soc Psych* 1992; **13**:9–24.
21 Sehnert S, Franks B, Yap AJ, Higgins ET. Scarcity, engagement, and value. *Motiv Emot* 2014; **38**:823–831.
22 Luck SJ, Woodman GF, Vogel EK. Event-related potential studies of attention. *Trends Cogn Sci* 2000; **4**:432–440.
23 Nobre AC, Sebestyen GN, Miniussi C. The dynamics of shifting visuospatial attention revealed by event-related potentials. *Neuropsychologia* 2000; **38**:964–974.
24 Kam JWY, Dao E, Blinn P, Krigolson OE, Boyd LA, Handy TC. Mind wandering and motor control: off-task thinking disrupts the online adjustment of behavior. *Front Hum Neurosci* 2012; **6**:329.
25 Rosch KS, Hawk LW Jr. The effects of performance-based rewards on neurophysiological correlates of stimulus, error, and feedback processing in children with ADHD. *Psychophysiology* 2013; **50**:1157–1173.

Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.