RESEARCH NOTE

Recalibration of perceived time across sensory modalities

**James V. M. Hanson · James Heron ·
David Whitaker**

Abstract When formulating an estimate of event time, the human sensory system has been shown to possess a degree of perceptual flexibility. Specifically, the perceived relative timing of auditory and visual stimuli is, to some extent, a product of recent experience. It has been suggested that this form of sensory recalibration may be peculiar to the audiovisual domain. Here we investigate how adaptation to sensory asynchrony influences the perceived temporal order of audiovisual, audiotactile and visuotactile stimulus pairs. Our data show that a brief period of repeated exposure to asynchrony in any of these sensory pairings results in marked changes in subsequent temporal order judgments: the point of perceived simultaneity shifts toward the level of adaptation asynchrony. We find that the size and nature of this shift is very similar in all three pairings and that sensitivity to asynchrony is unaffected by the adaptation process. In light of these findings we suggest that a single supramodal mechanism may be responsible for the observed recalibration of multisensory perceived time.

Keywords Temporal order judgment · Adaptation ·
Psychophysics · Multisensory processing ·
Perceived timing

Introduction

At least two physical factors influence the perceived timing of external multisensory events. Firstly, the differential velocities of light and sound in air ensure that distant auditory stimuli arrive progressively later than their visual counterparts. Secondly, neural transmission and conduction latencies ensure that visual, auditory and tactile stimuli can arrive synchronously at their respective receptor surfaces, yet arrive asynchronously at their respective primary cortices (Fain 2003). As such, it could be considered surprising that the auditory, visual and tactile components of a single event are usually perceived as simultaneous.

The fact that, under most real-world situations, perceptual synchrony is the norm itself suggests that the nervous system possesses a degree of temporal elasticity, which allows it to recalibrate signals in different sensory modalities to maintain a veridical percept of the world. In recent years, this concept of temporal sensory recalibration has been the subject of mounting scientific interest (Haggard et al. 2002; Sugita and Suzuki 2003; Fujisaki et al. 2004; Kopinska and Harris 2004; Lewald and Guski 2004; Vroomen et al. 2004; Arnold et al. 2005; Harrar and Harris 2005; Pesavento and Schlag 2006; Stetson et al. 2006; Keetels and Vroomen 2007; Navarra et al. 2007).

A novel mechanism for temporal sensory recalibration has been described by both Fujisaki et al. (2004) and Vroomen et al. (2004). Both studies demonstrated shifts in the point of subjective simultaneity (PSS—the physical temporal offset between two stimuli required for perceptual simultaneity) following exposure to a series of asynchronous audiovisual stimulus pairs. Specifically, following exposure to asynchronous pairs (e.g. vision leads sound), subjects required the same polarity of asynchrony for the two to be perceived as simultaneous (e.g. vision must now physically lead sound for perceptual simultaneity).

On the basis that only distant audiovisual events suffer the aforementioned temporal misalignment (vision physically leads sound at significant distances), it has been

J. V. M. Hanson (✉) · J. Heron · D. Whitaker
Department of Optometry, University of Bradford,
Richmond Road, Bradford BD7 1DP, UK
e-mail: j.v.m.hanson@bradford.ac.uk

suggested that this type of "lag adaptation" mechanism is likely to be restricted to the audiovisual domain (Miyazaki et al. 2006). By definition, a tactile stimulus has to be in direct contact with an observer. Thus, there appears to be no obvious ecological advantage in possessing a lag adaptation mechanism that involves this modality (Miyazaki et al. 2006). However, this argument fails to account for the earlier finding that recalibration readily occurs following exposure to "sound leads vision" (Fujisaki et al. 2004; Vroomen et al. 2004), despite the impossibility of such a situation in a natural environment. This suggests that temporal flexibility may not be directly related to long-term exposure to "vision leads sound", thereby introducing the possibility of adaptive timing shifts involving the tactile modality. To date, a systematic investigation of asynchrony adaptation in the three possible sensory pairings has not been undertaken. With this in mind, the present work aimed to test the hypothesis of Miyazaki et al. (2006) that temporal recalibration is unique to audiovisual stimulus pairings.

Methods

Observers were authors J.V.M.H., D.W. and J.H., and trained observer C.V. (who was naive as to the purpose of the experiment). All observers gave their informed consent to take part in the experiment, which was carried out in accordance with the 1964 Declaration of Helsinki. The stimuli were a 10 ms duration flash of an LED (10 mm diameter, luminous intensity 600 cd/m^2) viewed at a distance of 55 cm, a 10 ms square-wave windowed white noise burst (70 dB SPL) delivered binaurally over Sennheiser HD650 headphones, and a 10 ms tap on the left forefinger delivered via an electrical solenoid (whose sound output was rendered inaudible to observers). Presentation of the stimuli was controlled by custom-written software run in MatLab (Mathworks, USA) on a Dell desktop PC. The relative timings of all stimuli were verified by simultaneous capture on a multiple trace oscilloscope.

Procedures

Baseline measures of PSS were obtained for each observer in each potential modality pairing (audiovisual, audiotactile and visuotactile) by use of a temporal order judgment (TOJ) task. A trial consisted of pairs of stimuli presented at one of seven physical stimuli onset asynchronies (SOA): −90, −60, −30, 0 (simultaneous), 30, 60 and 90 ms, which were randomly interleaved within a method of constant stimuli. In the present study, positive SOAs always refer to

a visual lead and negative always to a tactile lead. The interstimulus interval (ISI) varied randomly (with a uniform probability) between 250 and 750 ms. After each presentation, the observer made an unspeeded TOJ as to "which modality came first" and responded via the computer keyboard. Each experimental run contained ten presentations at each SOA and each observer completed five experimental runs in each of the three stimulus pairings, giving a total of 50 repetitions per SOA, per observer, for each modality pairing.

In the adapted conditions, an initial period of adaptation preceded testing; this consisted of 100 pairs of stimuli separated by an SOA of either +90 or −90 ms. This value was chosen because previous work suggested that this asynchrony level was sufficient to elicit quantifiable adaptive shifts in PSS (Fujisaki et al. 2004; Vroomen et al. 2004). The time between presentations of stimulus pairs varied randomly (with a uniform probability) between 250 and 750 ms. Observers simply attended to these stimulus pairs without being required to make any judgment as to their temporal relationship. After the adaptation phase, a pause of 1-s duration alerted the observer that the adaptation phase was complete and the test phase was imminent.

This 1-s pause was followed by four "top-up" asynchronous (e.g. ±90 ms) stimulus pairs—identical to those in the adaptation phase—plus a fifth "test" pair with one of seven SOAs separated by 30 ms steps. As preliminary data showed significant PSS shifts in the adapted conditions, the range of test SOAs was offset in an appropriate direction so as to obtain a symmetrical psychometric function (in this case 30 ms). By making their TOJ response to this test pair (in an identical manner as in the baseline, unadapted condition), the next phase of the top-up/test cycle was initiated. Each experimental run consisted of ten test presentations at each SOA, and each observer completed five experimental runs following both positive and negative asynchrony adaptation (+90 and −90 ms) in each of the three modality pairings, making a total of 50 repetitions per SOA per observer for each condition tested.
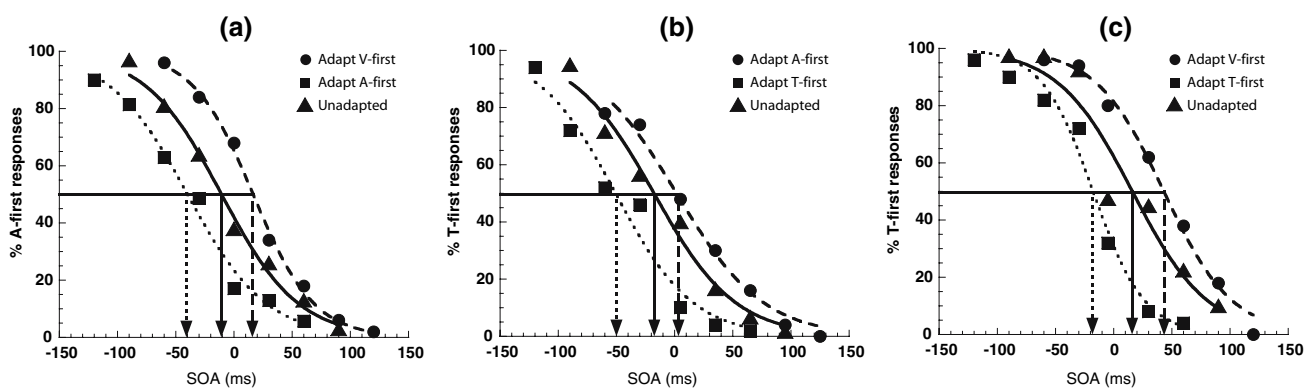
The percentage of "sound-first" or "touch-first" responses was plotted as a function of SOA and fitted with a logistic function of the form

$$y = \frac{100}{1 + e^{-\frac{(x-\mu)}{\theta}}}$$

where μ is the SOA equal to the PSS (corresponding to the 50% response level on the psychometric function), and θ provides an estimate of sensitivity to temporal order in the form of just-noticeable difference (JND) (approximately half the offset between the 27 and 73% response levels on the psychometric function).

Results

Psychometric functions for representative observer JVMH can be seen in Fig. 1 (a–c, corresponding to the three sensory pairings). Inspection of this figure reveals a clear lateral separation between the adapted functions (circles/long-dashed curve and squares/short-dashed curve) and their respective unadapted baseline functions (solid curve/triangles) for all sensory pairings. Following adaptation, perceived asynchrony shifts away from baseline toward the adaptation phase's asynchrony value. For example, adaptation to +90 ms of AV asynchrony (vision leads sound) increases the post-adaptation proportion of "sound-first" responses (Fig. 1a, filled circles) such that a visual lead is now required for perceived synchrony. The reverse pattern of results can be observed for the −90 ms AV adaptation condition (Fig. 1a, filled squares). For each function, arrows indicate the relevant PSS on the *x*-axis. Figure 2a shows the shift in PSS (relative to baseline) for all sensory pairings averaged across all four observers. The pattern of results is consistent for all conditions—perceived timing is strongly influenced by recent experience, irrespective of the type of sensory pairing or the polarity of the adapting asynchrony. Adapted PSS shifts for all observers were combined within a 2-way repeated measures ANOVA which revealed that the mean PSS shift was not significantly different across the three modality pairings ($F_{2,6} = 2.47$, $P > 0.05$), but that the polarity of adaptation within each modality pairing was highly significant ($F_{1,3} = 53.08$, $P = 0.0053$). There was no significant interaction between these two factors ($F_{2,6} = 2.13$, $P > 0.05$), indicating that the effect of polarity was consistent across each modality pairing.

Figure 2b shows average sensitivity to temporal order in the form of a just-noticeable difference (JND). A second 2-way repeated measures ANOVA was performed to analyse the JND data; this revealed that JNDs did not differ significantly across sensory pairings ($F_{2,6} = 0.93$, $P > 0.05$) and did not differ significantly between the three conditions (baseline and both polarities of asynchrony) ($F_{2,6} = 2.37$, $P > 0.05$). In other words, JND was found not to differ significantly before and after asynchrony adaptation.

Although the data presented above clearly indicate significant adaptive shifts in PSS following asynchrony adaptation in all three sensory pairings, a potential criticism is that observers could have unconsciously adopted a strategy of balancing their TOJs (e.g. 50% "vision first", 50% "sound first") during the method of constant stimuli, thereby forcing the PSS to be centred upon the mean of pre-chosen stimulus levels. All four observers therefore participated in an additional experiment designed to eliminate this possibility.

In this control experiment, baseline and post-adaptation PSS was obtained via the use of a PEST staircase (Taylor and Creelman 1967). The initial SOA was unknown to observers and could vary randomly between ±100 ms. Fifty stimulus pairs were presented during each experimental run, and PSS was calculated as the mean of the presentation offsets after the first three reversals of the staircase. The adaptation and top-up experimental phases were identical to the previous experiment, with each test presentation offset being unknown to the observer and dependent on the staircase procedure. PSS values for each observer were derived from three repetitions of the staircase procedure in each condition.

Fig. 1 Sample psychometric functions for representative observer JVMH for audiovisual (*AV*, **a**), audiotactile (*AT*, **b**) and visuotactile (*VT*, **c**) pairings. *Solid curves* and *triangles* are derived from unadapted TOJs in each pairing, which represent the "baseline" condition. *Long dashed curves* and *circles* are derived from TOJs made following adaptati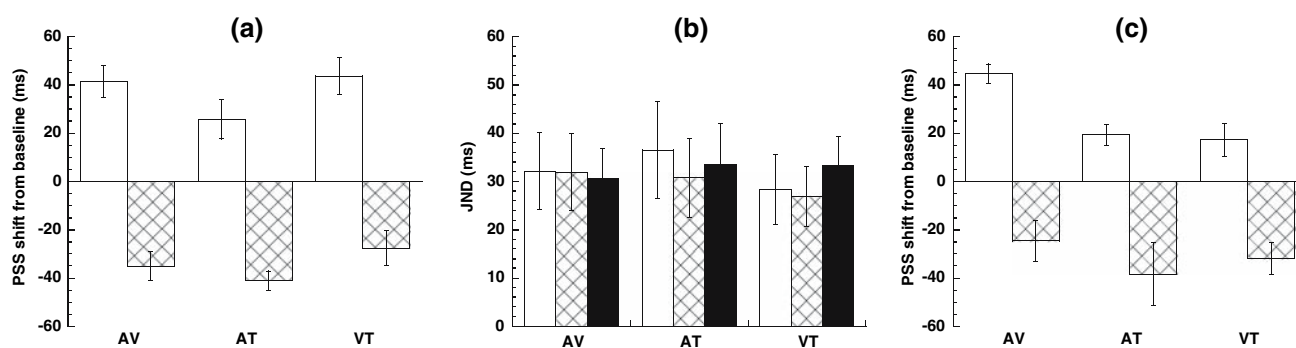on to positive asynchrony (positive always referring to a visual lead), whilst *short dashed curves* and *squares* are derived from TOJs made following adaptation to negative asynchrony (negative always referring to a tactile lead). The physical temporal offset between the stimuli which corresponds to chance performance is the point of subjective simultaneity (PSS). For each *curve*, *arrows* indicate the relevant PSS on the *x*-axis

Fig. 2 a Shows PSS shifts from baseline following asynchrony adaptation in the main experiment, averaged across four observers. The *unfilled bars* show shifts in PSS from baseline following adaptation to positive asynchrony (vision first in the AV and VT pairings, sound first in the AT condition); the *hatched bars* show shifts in PSS from baseline following adaptation to negative asynchrony (touch first in the AT and VT conditions, sound first in the AV condition). **b** Shows JND after adaptation to positive asynchrony (*unfilled bars*), negative asynchrony (*hatched bars*) and in the baseline condition (*solid filled bars*). The values are averaged across all four observers. **c** Shows PSS shifts from baseline following asynchrony adaptation in the control experiment; in this case all PSS values were obtained by the use of a PEST staircase procedure (see main text for details). The figure is otherwise identical to **a**. *Error bars* represent the SEM for each condition

The results of this control experiment are shown in Fig. 2c. A two-way repeated measures ANOVA revealed that, as before, the PSS shift was not significantly different across the sensory pairings ($F_{2,6} = 3.7$, $P > 0.05$) but there was a highly significant effect of adaptation polarity ($F_{1,3} = 100.7$, $P = 0.0021$) without any significant interaction effect ($F_{2,6} = 0.778$, $P > 0.05$). If observers simply balanced their TOJs from the start of the staircase, their final PSS values would faithfully reflect the original random asynchrony value (± 100 ms) that initiated the staircase. Figure 2c shows this is definitively not the case. Thus, the presence of a robust recalibration of perceived time for all sensory pairings is found to be consistent across different methodologies.

Discussion

The results of the current study demonstrate that perceived timing is markedly influenced by recent experience, irrespective of the nature of the sensory pairing. Whilst our audiovisual effects confirm previously published effects (Fujisaki et al. 2004; Vroomen et al. 2004) our visuotactile and audiotactile effects form the first convincing demonstration of adaptive temporal realignment involving the tactile domain. The robustness of the observed effect is demonstrated by the close correspondence between the results of the main experiment and the control experiment, which removed a potential confound in the form of the offset SOA range by the use of a PEST staircase procedure.

A recent study by Navarra et al. (2007) investigated a single asynchrony polarity for a single sensory pairing. They found that adaptation to 75 ms "sound leads touch" induced a small (12 ms) increase in thresholds relative to their "adapt synchronous" condition without any significant change in PSS. This is the exact opposite finding to ours and therefore requires consideration. Numerous methodological differences between the two studies make comparison problematic. For example, the magnitudes of our adaptive PSS shifts are judged relative to an unadapted baseline whereas Navarra et al. (2007) use a physically synchronous audiotactile pairing as their baseline against which to compare their adapted results. Given the well-documented dissociation between physical and perceived synchrony in the audiotactile domain (Zampini et al. 2005; Navarra et al. 2007) a physically synchronous audiotactile-adapting stimulus may or may not be perceptually synchronous (e.g. Fig. 1b shows this observer has an audiotactile baseline PSS of -17 ms). However this factor (in isolation) is unlikely to provide a complete explanation for the dissociation between our data and that of Navarra et al. (2007). Another significant difference is Navarra et al.'s alteration of stimulus characteristics between adaptation and test phase. If adaptation modulates perceived timing at a relatively early processing stage, the manipulation of low-level stimulus characteristics may prevent transfer of PSS shifts between the two experimental phases. However, caution must be applied when making assumptions about the neural locus of asynchrony adaptation: our data suggest a relatively late-stage mechanism that recalibrates any sensory input with respect to recent experience. Specifically, the size of our effects is broadly comparable across modality pairings suggesting that the recalibration mechanism may modulate the same fixed proportion of the adaptation asynchrony (approximately 30% on average) regardless of the modality of the signal itself.

In the present study, asynchrony JNDs were found to be invariant across modality pairings and adapted versus

baseline conditions. As such, it appears the nervous system has the ability to adaptively recalibrate sensory temporal relationships without a discernable loss of sensitivity. This agrees with anecdotal reports from observers who felt that by the end of the adaptation phase the physically asynchronous stimulus pairs felt markedly closer to being perceptually synchronous. The JND data suggest this phenomenon is not a product of a progressive loss in sensitivity—rather, the signals (relative to one another) are subject to an adaptive temporal recalibration. This perceptual recalibration is consistent with Helson's (1964) adaptation-level theory. In Helson's framework, human sensations are judged relative to an aggregate of recent experience which is constantly updated by novel sensory stimulation (Behar and Bevan 1961). In the case of the present work, it seems likely that the fixed adapting asynchrony provides our observers with a revised perceptual "anchor" (e.g., sound leads vision) around which to centre their adaptation level.

Returning to the hypothesis of Miyazaki et al. (2006), our adaptive shifts argue against the idea that the effects observed here and by others (Fujisaki et al. 2004; Vroomen et al. 2004; Heron et al. 2007) are peculiar to the audiovisual domain. Miyazaki et al. (2006) speculate that if the perceptual consequences of repeated exposure to tactile signals are to be explained within a Bayesian framework, subjects will impose perceptual asynchrony onto physically simultaneous signals. This asynchrony will not be a classic "rebound" effect; rather, it will be similar to that of the adapting stimulus. In other words, our psychometric functions would be laterally displaced in the opposite direction to that shown in Fig. 1. Both Bayesian and "lag adaptation" (as Miyazaki terms the "rebound" effects reported here) may be active, competing mechanisms, with the perceptual outcome dependent on which modalities are tested (Miyazaki et al. 2006). In the case of audiovisual asynchrony, the "lag adaptation" mechanism has been proposed to overcome the magnitude of any Bayesian recalibration, whereas when tactile signals are involved Bayesian-type recalibration dominates. The latter outcome has been attributed to observers adopting a new prior assumption of tactile temporal order. This prior corresponds to the asynchrony to which observers are repeatedly exposed to during the adaptation phase (Miyazaki et al. 2006). This concept has received support from a recent visuotactile study showing that new priors can be adopted by observers given just a single hour of training (Ernst 2007).

Alternatively, it is possible that the results of both studies may be compatible with Bayesian models of perception. Given that approximately simultaneous taps to one hand followed by the other (as employed by Miyazaki et al. (2006)) are a relatively infrequent "real-world" occurrence, observers are unlikely to possess a deeply ingrained prior preference for their simultaneity. In this situation, observers may be more likely to modify any pre-existing prior assumptions about the temporal relationship between the signals (Ernst 2007). In contrast, it is reasonable to speculate that a lifetime's experience of close temporal correlation may produce a strong prior assumption of synchrony. As mentioned in the introduction, this would be the case for both proximal audiovisual events and all audiotactile and visuotactile events whose sensory components must—by definition—occur simultaneously and arrive at their respective receptor surfaces approximately simultaneously. Post-adaptation "rebound" effects of the type described here could well be the product of a tendency to realign perception in a manner that maintains concordance with prior assumptions. A recent study has proposed that adaptation may alter the likelihood function (Mamassian et al. 2002) without influencing any pre-existing priors (Stocker and Simoncelli 2005). This implies that when observers expect synchronous sensory signals they realign their sensory timing rather than their pre-existing knowledge about their environment.

In conclusion, the results of the current study demonstrate a degree of temporal flexibility that appears comparable across the range of sensory pairings. This cannot be explained simply in terms of maintaining veridicality in the time domain. Rather, it seems that the nervous system may realign any temporal signal that fails to meet its prior assumptions about the world. Of course, the task of formulating a sensory estimate of "when" is far from abstract—our motor commands are only as accurate as the sensory estimates that guide them. A critical question for future work is how purely sensory adaptation (as described in the current study) relates to recent examples of sensorimotor temporal recalibration (Pesavento and Schlag 2006; Stetson et al. 2006).

References

Arnold DH, Johnston A, Nishida S (2005) Timing sight and sound. Vis Res 45:1275–1284

Behar I, Bevan W (1961) The perceived duration of auditory and visual intervals: cross-modal comparison and interaction. Am J Psychol 74:17–26

Ernst MO (2007) Learning to integrate arbitrary signals from vision and touch. J Vis 7:1–14

Fain GL (2003) Sensory transduction. Sinauer Associates, Sunderland

Fujisaki W, Shimojo S, Kashino M, Nishida S (2004) Recalibration of audiovisual simultaneity. Nat Neurosci 7:773–778

Haggard P, Clark S, Kalogeras J (2002) Voluntary action and conscious awareness. Nat Neurosci 5:382–385

Harrar V, Harris LR (2005) Simultaneity constancy: detecting events with touch and vision. Exp Brain Res 166:465–473

Helson H (1964) Adaptation-level theory. Harper & Row, New York

Heron J, Whitaker D, McGraw PV, Horoshenkov KV (2007) Adaptation minimizes distance-related audiovisual delays. J Vis 7:1–8

Keetels M, Vroomen J (2007) No effect of auditory–visual spatial disparity on temporal recalibration. Exp Brain Res 182:559–565

Kopinska A, Harris LR (2004) Simultaneity constancy. Perception 33:1049–1060

Lewald J, Guski R (2004) Auditory–visual temporal integration as a function of distance: no compensation for sound-transmission time in human perception. Neurosci Lett 357:119–122

Mamassian P, Landy M, Maloney LT (2002) Bayesian modelling of visual perception. In: Rao RPN, Olshausen BA, Lewicki MS (eds) Probabilistic models of the brain: perception and neural function. MIT Press, Cambridge, pp 13–36

Miyazaki M, Yamamoto S, Uchida S, Kitazawa S (2006) Bayesian calibration of simultaneity in tactile temporal order judgment. Nat Neurosci 9:875–877

Navarra J, Soto-Faraco S, Spence C (2007) Adaptation to audiotactile asynchrony. Neurosci Lett 413:72–76

Pesavento MJ, Schlag J (2006) Transfer of learned perception of sensorimotor simultaneity. Exp Brain Res 174:435–442

Stetson C, Cui X, Montague PR, Eagleman DM (2006) Motor-sensory recalibration leads to an illusory reversal of action and sensation. Neuron 51:651–659

Stocker AA, Simoncelli EP (2005) Sensory adaptation within a Bayesian framework for perception. In: Weiss Y, Scholkopf B, Platt J (eds) Neural information processing systems (NIPS), vol 18. MIT Press, Vancouver

Sugita Y, Suzuki Y (2003) Audiovisual perception—implicit estimation of sound-arrival time. Nature 421:911–911

Taylor MM, Creelman CD (1967) PEST: efficient estimates on probability functions. J Acoust Soc Am 41:782–787

Vroomen J, Keetels M, de Gelder B, Bertelson P (2004) Recalibration of temporal order perception by exposure to audio–visual asynchrony. Cogn Brain Res 22:32–35

Zampini M, Brown T, Shore DI, Maravita A, Roder B, Spence C (2005) Audiotactile temporal order judgments. Acta Psychol 118:277–291