Illuminating ATOM: Taking time across the colour category border

Samuel, Steven[a, b]

Bylund, Emanuel[d]

Cooper, Rachel[b]

Athanasopoulos, Panos[c]

[a] Stockholm University, Sweden.

[b] University of Essex, U.K.

[c]Lancaster University, U.K.

[d]Dept. Of General Linguistics, Stellenbosch University, Sweden.

Accepted for publication in Acta Psychologica:

doi.org/10.1016/j.actpsy.2018.01.011

Word count: 7299 (excluding abstract, acknowledgements, references and tables, including footnotes, captions and appendix).

Key words: magnitude, colour categories, duration perception, cross-linguistic, Whorf

Declaration of interest: none

Address for correspondence: Address for correspondence: Steven Samuel. University of Cambridge, Department of Psychology, Downing Street, CB2 3EB. Tel: 01223 333563. Email: ss2391@cam.ac.uk

Abstract

Walsh's A Theory Of Magnitude (ATOM) contends that we represent magnitudes such as number, space, time and luminance on a shared metric, such that 'more' of one leads to the perception of 'more' of the other (e.g. Walsh, 2003). In support of ATOM, participants have been shown to judge intervals between stimuli that are more discrepant in luminance as having a longer duration than intervals between stimuli whose luminance differs by a smaller degree (Xuan, Zhang, He, & Chen, 2007). We tested the potential limits to the ability of luminance to influence duration perception by investigating the possibility that the luminance-duration relationship might be interrupted by a concurrent change in the colour of that luminance. We showed native Greek and native English speakers sequences of stimuli that could be either light or dark versions of green or blue. Whereas for both groups a shift in green luminance does not comprise a categorical shift in colour, for Greek speakers shifts between light and dark blue cross a colour category boundary (*ghalazio* and *ble* respectively). We found that duration judgements were neither interrupted nor inflated by a shift in colour category. These results represent the first evidence that the influence of luminance change on duration perception is resistant to interference from discrete changes within the same perceptual input.

1. Introduction

A sense of time passing is a cornerstone of the human experience, yet unlike vision, hearing and such we have no ‘organ’ for it (Ornstein, 1969). Instead, when we subjectively measure ‘duration’, we are not really sampling objective time but rather measuring perceptual *events* (Pöppel, 2009). Research has made it abundantly clear that our subjective perception of duration is fallible, and even influenced in systematic fashion, by factors external to time itself (see Bueti & Walsh, 2009; Droit-Volet & Gil, 2009, for reviews). For example, clusters of more numerous stimuli are perceived to last longer than fewer stimuli (Alards-Tomalin, Walker, Kravetz, & Leboe-McGowan, 2015; Dormal, Seron, & Pesenti, 2006; Xuan, Zhang, He, & Chen, 2007), changing stimuli longer than unchanging stimuli (Poynter & Homa, 1983), larger numbers longer than smaller numbers (Alards-Tomalin et al., 2015; Oliveri et al., 2008; Vicario, Pecoraro, Turriziani, Koch, Caltagirone, & Oliveri, 2008; Xuan, et al., 2007), and brighter stimuli longer than dimmer stimuli (Xuan et al., 2007).

One account for such findings is Walsh’s influential ‘A Theory of Magnitude’ (ATOM; for reviews see Bueti & Walsh, 2009; Walsh, 2003; Winter, Marghetis, & Matlock, 2015). ATOM proposes that there is a shared metric for representing magnitudes of time, space and number, as well as other codes such as luminance. This shared metric likely arose from the need to control actions. For example, throwing and catching require a judgment to be made about the size of the object (number/quantity) and the object’s velocity (an equation formed through spatial and temporal variables). Crucially, an effect of this shared metric is that the perception of ‘more’ in one domain, such as number, leads to the perception of ‘more’ in another, such as duration. This overlap of magnitudes was then ‘scaled up’ to expand to other dimensions such as luminance through knowledge of number.

Crucially, ATOM designates specifically *analogue* format to the shared magnitude metric. The present experiment was designed to test the level of priority that the analogue metric proposed by ATOM actually receives. One test of this hypothesis is to see if this metric continues to process analogue information in the same way when that analogue change coincides with a simultaneous but non-analogue (i.e. discrete) change. Would changes in luminance continue to bias duration perception if the two stimuli that differed in luminance also differed in their *colour category*—a discrete- rather than analogue-level change—compared to a matched luminance change but without a colour change? Or would a shared magnitude processor cease to function in the same way, indicating that other, non-analogue features can have an impact?

1.1 The Kappa effect and the effect of luminance on duration.

The effect of luminance on time perception is not limited to the duration of a stimulus but extends to the perception of the *intervals* between stimuli. Previous studies have shown that participants' perception of the duration of the two intervals (AX and XB) in a three-stimulus 'AXB' sequence is influenced by the physical distance between the stimuli in question, such that if the AXB stimuli are light flashes and stimuli A and X are separated by a greater physical distance than X and B, the AX interval is more likely to be judged as temporally 'long' than XB, all else being equal. This perceptual illusion, known as the Kappa effect (e.g. Abe, 1935; Cohen, Hansel, & Sylvester, 1954; Sarrazin, Giraudo, Pailhouse, & Bootsma, 2004), has been extended to differences other than physical distance between stimuli. Xuan et al. (2007) found that participants were influenced by the luminance of stimuli when judging whether the intervals between them were long or short in duration, such that more errors were made when the temporally longer interval was bounded by a relatively subtle change in luminance compared to the temporally shorter interval, and vice-versa. Similar results were found when the magnitude in question was not luminance but numerosity (number of dots in the display), number magnitude (Arabic numbers) and size (varying the size of squares). Since Xuan et al.'s study, the subjective 'warping' of time through the magnitudes of size and numerals has been replicated (Alards-Tomalin, Leboe-McGowan, Shaw, & Leboe-McGowan, 2014), and has been extended to magnitudes of auditory intensity (intervals bounded by tones that were closer in intensity were perceived to be shorter: Alards-Tomalin, Leboe-McGowan, & Mondor, 2013), and even levels of colour saturation (Alards-Tomalin et al., 2014). Crucially, in all of these studies, participants knew that the properties of the stimuli were irrelevant to the task, but seemingly could not ignore the discrepancies in magnitude in these non-temporal dimensions. Xuan and colleagues suggested that their results support the ATOM proposal, whereas Alards-Tomalin et al. (2014) offered an alternative possibility; namely that stimuli that are perceived to be more proximal (e.g. 1 and 2) are more likely to be integrated for the purposes of efficient processing, and hence the intervals bounded by these integrated stimuli may be subjectively compressed.

1.2. Language and duration perception.

Categorical Perception (CP) occurs when a continuum is perceived to alter in a discrete and qualitative way at a category boundary (Harnad, 1987, cited in Roberson, Pak, & Hanley, 2008), and as such colours that fall within a category are perceived as more similar than colours from different categories, effectively 'distorting' a continuum (Roberson, Davidoff, Davies, & Shapiro, 2005). What in

English would be considered ‘blue’ colour space is ‘split’ into two categories in Greek, *ble* (approximately equivalent to ‘dark blue’ in English), and *ghalazio* (approximately ‘light blue’) (Androulaki, Gômez-Pestaña, Mitsakis, Jover, Coventry, & Davies, 2006; Athanasopoulos, 2009). A blue stimulus that changes from light to dark or vice-versa therefore crosses a category boundary in Greek, inducing CP, but not in English. Crucially for the present experiment, this means that a change in the luminance of two blue stimuli could be manipulated to cross a category boundary for one group—Greek speakers—but not the other. As such, by presenting Greek speakers with changes in ‘blue’ luminance we could introduce a discrete change in the identity of that stimulus to coincide with the analogue change in luminance, but we could show precisely the same stimuli to English speakers and now the change would be perceived as belonging to one colour alone.

Importantly, the *ble*/*ghalazio* distinction has also already been shown to produce CP in Greek speakers. Thierry, Athanasopoulos, Wiggett, Dering, and Kuipers (2009) recorded ERP data from Greek and English speakers while they watched sequences of light and dark blue and light and dark green stimuli. Participants were instructed to make occasional responses based on shape, and not their colour. Using visual mismatch negativity (vMMN) as a measure of pre-attentive change perception, Thierry and colleagues found larger vMMN for changes in blues than greens in the Greek speakers, but equivalent performance across both colours in the native English speakers. The researchers concluded that the results demonstrated an unconscious effect of linguistic categories on colour perception.

Support for CP effects of colour terms comes also from behavioural experiments between speakers of other languages who show a similar division of the blue colour space into two separate basic terms. It has been shown that speakers of

Russian, which also divides the 'blue' colour space into two basic terms marking similar colour spaces to those in Greek, are faster to visually locate an 'oddball' light blue stimulus among a set of dark blues (or vice versa) than speakers of English (Winawer, Witthoft, Frank, Wu, Wade, & Boroditsky, 2007). Similar findings have been reported in Korean speakers, who distinguish between what in English would be termed 'yellow-green' and 'green' (Roberson et al., 2008). In both of these studies these results have proven to be strongest in, or unique to, presentation of stimuli in the right visual field and hence in receipt of privileged processing by the language regions of the brain in the left hemisphere, adding further weight to the argument that the colour CP effect has its source in acquired linguistic categories.

1.3. The present experiment

Xuan and colleagues (2007) looked at the effect of luminance on duration judgments in a population that views the changes in luminance as a continuum. We here test the robustness of ATOM's predictions when language imposes a categorical boundary on the luminance dimension. Using a similar design to that adopted by Alards-Tomalin and colleagues (2014), we presented native Greek and native English speakers with AXB sequences of blue or green colour patches. Crucially, either the first (A) or last (B) patch in the sequence consisted of a 'dark' or 'light' version of the colour found in the other two stimuli in the triplet, meaning that one interval was always bounded by stimuli that changed in luminance, and one interval by stimuli that were identical. When these stimuli were blue, the change in luminance also introduced a categorical shift (a CP effect) in the colour term of that stimulus for the Greek speakers, but not for the English speakers. We hypothesised that, if luminance and duration magnitude share processing through a single, analogue metric, the effect of luminance on duration should proceed regardless of the simultaneous occurrence of a discrete change in colour in the luminance input stream. Alternatively, if CP interferes in some way with the processing of the analogue luminance change, then we should see a difference in judgments of duration compared to trials where no CP was introduced.

If CP effects do influence duration perception, the way in which they do so would also be informative as to the underlying processes that lead to the influence of magnitude on duration perception. Specifically, an *underestimation* of duration when CP occurs should arise if the *ble/ghalazio* interval causes the processing of luminance to be *halted* at the category border. Such a result would indicate an exception to the notion of luminance influencing duration, and impose a new constraint on the shared metric account. Alternatively, an *overestimation* of duration is what might be predicted if it is harder to integrate stimuli that come under different linguistic categories. Such a pattern would support the proposal that it is the integration of proximal stimuli that leads to temporal compression effects (e.g. Alards-Tomalin et al., 2014).

We applied a number of controls to our design. Firstly, for any such effects to be truly the result of CP, they should occur *only* within the Greek-speaking group and not occur simultaneously in native English speakers, for whom the changes in blue luminance to not constitute a shift in colour category. Simultaneously, CP effects should occur only for *blue* stimuli, and not for green stimuli, for which both the English and Greek speakers use one basic category (‘*prasino’* in Greek). Although we chose to recruit Greek speakers in the UK, it is well-known that a bilingual’s first language is always active (e.g. Marian & Spivey, 2003a, 2003b; Thierry & Wu, 2007; Wu & Thierry, 2010), and Greek speakers recruited in the UK (and after an average

18 months in the country) continue to show sensitivity to the *ble/ghalazio* boundary (Thierry et al., 2009). Recruiting participants of both languages from the same university allowed us to control for environmental factors that might have otherwise influenced our findings. However, since the category boundaries and focal colours of *ble* and *ghalazio* have been shown to shift with increased acculturation to an English-language environment (Athanasopoulos, 2009; Athanasopoulos, Dering, Wiggett, Kuipers, & Thierry, 2010), we recruited twice as many native Greek speakers as native English speakers with a view to splitting the former into short- and long-stay groups and comparing these two groups to the native English speakers. This allowed us to test whether any effects of the *ble*/*ghalazio* boundary might be stronger in (or limited to) those native Greek speakers who had spent only a very short time in an English-speaking environment. Finally, we also included a block with symbolic number stimuli to verify that the ‘classic’ number/time link was also true for our sample of Greek and English speakers, and to test whether one group was perhaps intrinsically more susceptible to effects magnitude change on duration perception, such that any effects that we may have attributed to the crossing of a colour category boundary might instead be attributable to greater sensitivity to magnitude in Greek speakers than English speakers more generally.

2. Method

2.1. Participants and colour naming tasks

We recruited 89 participants from the same UK university, of which 62 were native speakers of Greek and 28 native speakers of English. All were compensated financially for their time. Four failed to follow instructions and their data were discarded (3 Greek speakers, 1 English speaker). Four (1 English, 3 Greek) were excluded for extreme outlying responses, one English speaker for evidence of a potential colour vision deficit, and two Greek speakers due to technical faults. All had normal or corrected-to-normal vision and colour vision (score of 15 or 16 out of 16 on the City University Colour Vision Test, 3rd edition, 1998).

All the native Greek speakers bar one, whose data were excluded, named the light blue colour *ghalazio* and the dark blue colour *ble* in a simple forced choice task in which each blue was presented side-by-side on the screen. In addition to this task, all participants (native Greek and native English speakers) also performed a free-choice naming task, in which they saw patches of the colours used presented in a fixed order (Light Green, Dark Blue, Dark Green, Light Blue) and were asked to name the colours in their first language. This task was performed as part of another experiment that was always completed on a different day, prior to the experiment described here. We included the free-choice task because it allowed us to examine whether any effects of crossing a *ble/ghalazio* boundary were modulated by whether these stimuli were also freely and preferentially associated with these terms rather than with similar, perhaps more literary alternatives (e.g. Greek equivalents of 'sky' or 'cerulean' for light blue). It is because we anticipated splitting the native Greek speakers into two groups not just on the basis of their length of stay (to check for any effects of acculturation), but also on the basis of their natural adherence to the *ble/ghalazio* terms for the stimuli, that we collected data from a sample of native English speakers of approximately half the size of the full Greek sample.

Final group numbers were 20 native English speakers and 50 native Greek speakers. The Greek speakers (M_{Age} = 23, range 18-33; 35 females) continued to use Greek on average 45% of the time and had spent on average 22 months in the UK (range 1-108months) at the time of testing. They typically began learning English at the age of 9, and had a good level of English by self-report (M = 4.28/5, range 3-5). These Greek speakers resembled those in the study by Thierry and colleagues (2009); their sample of native Greek speakers had spent 18 months in the UK at the time of testing (they were also recruited from a UK university) and displayed sensitivity to the *ble/ghalazi*o boundary.

The English speakers (M_{Age} = 28, range 18-47; 9 females) reported no knowledge of a second language that divided the light/dark blue border. All the English speakers used the words ‘green’ and ‘blue’ as head nouns for the relevant stimuli in the free choice naming task (5 who did not were discarded). All the native Greek speakers used *prasino* or related secondary terms which can be defined as types of *prasino*— *lachani* (‘lettuce’ n = 6) or *kyparisi* (‘cypress’ n = 1)— to describe the greens[1] (5 who did not were discarded). Of these native Greek speakers, 19 also named the relevant blues *ble* and *ghalazio* in the free choice task. These Greek speakers were also allocated to a ‘Strong Category’ sub-group. The remaining 31 Greek speakers named the blue stimuli in the free choice task using different terms, such as *ble* (n = 12) and *mov* (‘mauve’, n = 7) for the light blue stimulus, and *mov* (n = 3) for the dark blue stimulus. These participants were allocated to a ‘Weak Category’ subgroup. Comparing the results of these two groups and the native English speakers would allow us to establish whether any effects of the *ble/ghalazio* boundary

[1] These terms are types of the basic category *prasino* in the same sense that ‘emerald’ or ‘lime’ are types of the basic category ‘green’. These terms are hence subsumed *within* the basic category terms *prasino*/green; they do not constitute a categorical shift in the way that *ble* and *ghalazio* do.

are strongest in, or limited to, those participants for whom the stimuli were particularly strongly associated with the labels in question.

2.2. Materials and procedure

All stimuli (colour patches and numerals) were presented using E-Prime 2.0 software and a 32-bit colour and Nvidia graphics card on a 20" Mitsubishi Diamond Pro 2070 set to an 85Hz refresh rate with 1024x768 resolution. The monitor was switched on at least 30 minutes before testing for normal operating temperature to be achieved. Participants sat in darkness measured at less than .01 cd/m^2 (measured with a Minolta LS-100) in the absence of light from the monitor, and used a centrally-placed chinrest which maintained a constant distance from the screen of approx. 60cm. An enforced ten-minute dark adaptation period with only the light of the monitor (grey screen) preceded the practice trials. A grey background was maintained throughout the experiment.

2.3. Duration judgment task.

Figure 1 displays the slides and possible AXB sequences used in the experiment.

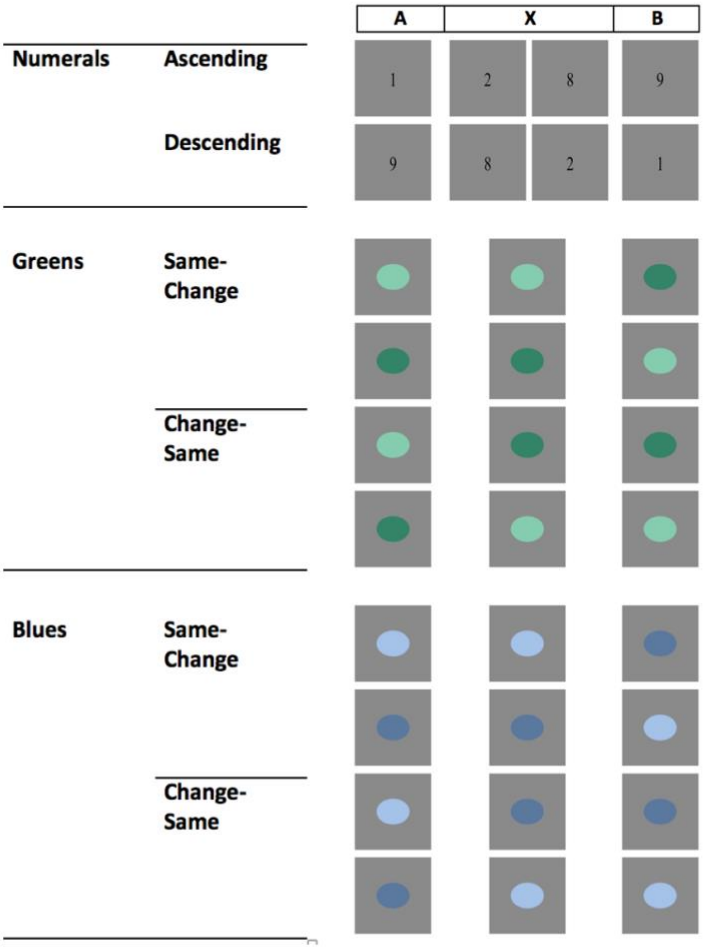

Figure 1: All possible AXB sequences for numbers and colours. For the numerals, the central ‘X’ stimulus was either 2 or 8, the choice of which created magnitude discrepancies of 1,7 or 7,1. For the blocks of colours, the luminance pattern could be either ‘same-change’ (first two slides identical, third a different luminance) or change-same (last two slides identical, first a different luminance). In the block of blue stimuli, a luminance change consisted of a colour category change for the Greek speakers only.

Colours. Following Alards-Tomalin et al. (2014), each patch was ovular and subtended a visual angle of 17.4 x 10.9. Observed *Yxy* values, measured with a PhotoResearch Spectrascan PR670, are detailed in the Appendix. The colour stimuli were generated from the same RGB co-ordinates as those used for light and dark blues and greens in Athanasopoulos et al. (2010) and Thierry et al. (2009). These colours had been selected for their cross-colour equivalence in luminance, their prototypicality as examples of the Greek categories for *ble* and *ghalazio,* and for the previous evidence that bilingual Greek-English speakers resident in the UK were sensitive to the distinction between them (Thierry et al., 2009).

Numerals. Stimuli were presented centrally in Times New Roman font and subtended a visual angle of 3.1^{o} x 5.7^{o}. The numerals used were ‘1’, ‘2’, ‘8’ and ‘9’, and were presented in either ascending order (i.e. AXB = 1,2/8, 9) or descending order (i.e. AXB = 9, 2/8, 1), forming magnitude patterns (1-7 and 7-1). Trials were equally divided between ascending and descending orders, 1-7 and 7-1 magnitude patterns, and the eight interval differences.

Trial procedure. Upon pressing the space bar, participants saw a fixation cross (500ms) before the three stimuli (AXB, each 200ms). As in Alards-Tomalin et al.’s (2014) experiment, intervals were calculated from the onset of the first slide (A) to the onset of the second (X), and the onset of the second (X) to the onset of the third (B). There were four ‘long-short’ interval duration patterns (785ms-635ms, 768ms-651ms, 752ms-668ms, 735ms-685ms), and four ‘short-long’ (635ms-785ms, 651ms-768ms, 668ms-752ms, 685ms-735ms). These eight patterns can be expressed in terms of the differences between the two interval durations (from short-long to long-short: -150ms, -117ms, -84ms, -50ms, 50ms, 84ms, 117ms, 150ms). Total trial time excluding the fixation cross was always 1420ms. After viewing each AXB sequence participants were prompted by the command ‘press’ to judge whether the time elapsed between the first and second slides (AX) was longer or shorter than the duration between the second and third slides (XB). Participants pressed ‘L’ on the keyboard to indicate they perceived a ‘Long-Short’ pattern and ‘S’ for ‘Short-Long’. Though no time limit was imposed, the instructions were to do this as quickly and accurately as possible, and to ignore the stimuli themselves making their choice.

Order of presentation. Numerals (N), greens (G) and blues (B) were presented in a blocked design with counterbalanced order (Of the 70 participants remaining following exclusions, 33% began with numerals, 30% with greens and 37% with blues), with 160 trials per block. Prior to beginning the task proper, participants completed a short practice session of 6 AXB sequences, all using the same stimuli that they were to see in the first experimental block, and all using the maximal interval duration difference of (+/-150ms). Feedback, both in terms of accuracy and speed of response, was provided during this practice, but never during the experimental blocks.

2.4. Analyses

Following the procedure used by Alards-Tomalin et al. (2014), we took the percentage of long-short responses as our dependent measure, which we analysed using repeated measures ANOVAs. In all cases, the data for each cell in these analyses were normally distributed (*p*s > .05 by Shapiro-Wilks tests) unless specifically highlighted in the text. There were no reported violations of the assumption of homogeneity in any test. Where Bayesian analyses were also conducted, we interpret the results in terms of the convention that meaningful evidence for a null result (i.e. a null that is sufficiently sensitive to suggest a meaningful test of the hypothesis rather than a null that is likely to be an artefact of low statistical power) requires a Bayes Factor (BF_{10}) of <0.33 (Dienes, 2014).

3. Results

The mean percentage of long-short responses for each interval and trial type for the entire sample (*N* = 70) is displayed in Table 1, and shows the expected pattern of a linear decline as the first interval becomes shorter relative to the second. This made clear that participants were indeed engaged in the task. Since further analyses would not yield any relevant information for our hypotheses concerning luminance and colour, we thus proceeded with our main analyses collapsing over these eight intervals.

Table 1. Percentage of long-short responses by interval (positive numbers indicate first interval was longer than the second) and trial type.

	Colours				Numerals	
	Same-Change		Change-Same		Small-Big	Big-Small
Interval	Green	Blue	Green	Blue		
+150ms	*64*	*66*	*66*	*68*	*67*	*67*
+117ms	*64*	*60*	*64*	*65*	*60*	*64*
+84ms	*56*	*57*	*64*	*63*	*55*	*60*
+50ms	*52*	*54*	*58*	*57*	*54*	*54*
-50ms	*38*	*38*	*48*	*48*	*41*	*48*
-84ms	*38*	*36*	*43*	*40*	*38*	*42*
-117ms	*33*	*30*	*39*	*39*	*35*	*36*
-150ms	*30*	*26*	*38*	*31*	*34*	*34*

3.1 Does crossing a colour category influence duration perception between blues and greens in Greek speakers?

We first analysed the data from the full complement of Greek speakers (*n* = 50). Each of these participants named the light blue stimulus *ghalazio* and the dark blue stimulus *ble* in the two-choice test.

We conducted a 2: Colour (Green vs. Blue) x 2: Luminance (Same-Change vs. Change-Same) repeated-measures ANOVA. The results are displayed in panel A of

Figure 2. As predicted, the analysis revealed a main effect of Luminance, $F(1, 49) = 11.532$, $MSE = 1.5$, $p = .001$, $\eta_p^2 = .191$, with more long-short responses given when the first interval coincided with a change in luminance (*M* = 53.3%, 95% CI [50.3, 56.2]), than when no change in luminance was perceived (*M* = 47.4%, 95% CI [44.4, 50.4]). There was no main effect of Colour, $F(1, 49) = 0.098$, $MSE = 1.0$, $p = .756$, $\eta_p^2 = .002$; participants gave a similar proportion of long-short responses for both greens (*M* = 50.6%, 95% CI [47.4, 53.7]) and blues (*M* = 50.1%, 95% CI [47.7,52.5]). Crucially, there was no interaction, $F(1, 49) = 0.108$, $MSE = 0.7$, $p = .744$, $\eta_p^2 = .002$. Native Greek speakers perceived the interval between stimuli that changed in luminance to be longer than the interval between stimuli that did not change, but there was no evidence that a change in colour category modulated this effect. To test the strength of this null interaction, we applied Bayesian analyses to the data. Crucially, the analysis found that the data were approximately 5 times more likely under the null hypothesis that there was no Colour x Luminance interaction, $BF_{10} = 0.211$, supporting the interpretation of a meaningful null according to Dienes (2014) criterion.

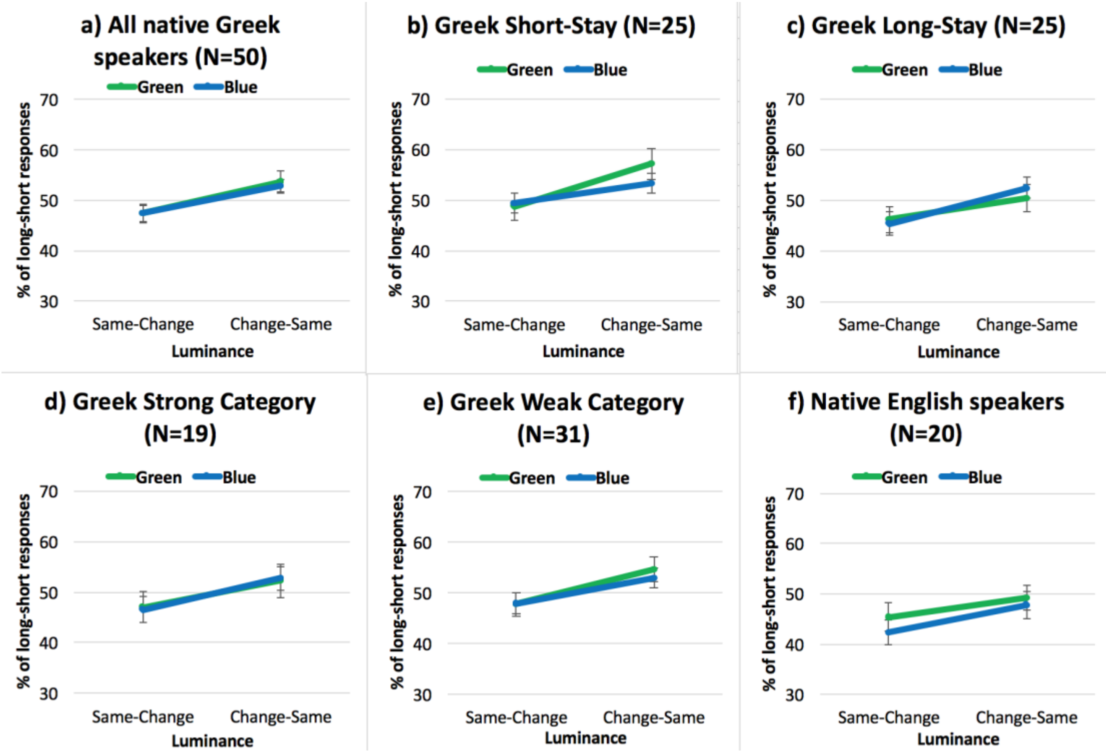

Figure 2. Results of analyses for colour stimuli (with standard error bars)

3.2 Are effects of the ble/ghalazio *boundary restricted to short-stay Greek speakers?*

If the total sample of 50 Greek speakers failed to show any influence of crossing a colour category boundary when making duration judgments, this could be because some had spent sufficient time in the UK for this boundary to become blurred. To test this possibility, we allocated the native Greek speakers into short- and long-stay groups by median split, and added the native English speakers as a third group in the analysis. The 'Long-Stay' group (n = 25) had spent an average of 40 months in the UK at the time of testing (95% CI [28,50]). By contrast, the 'Short-Stay' group (n = 25) had spent an average of only 4 months in the UK (95% CI [2,6]). We reasoned that if the *ble/ghalazio* border was more influential in those participants who had spent the least time in an English-language environment, then the effect of crossing the *ble/ghalazio* border should be different in the short-stay group relative to the long-stay group and the native English speakers, as well as relative to matched within-category changes in greens. Evidence for such an effect would thus be supported by a significant three-way interaction between Group (English vs. Greek Short-Stay vs. Greek Long-Stay), Colour (Green vs. Blue), and Luminance (Same-Change vs. Same-Change).

With the sole exception of green same-change trials in the Greek Short-Stay group (p = .022), the data did not deviate from normality, and hence we opted to proceed with parametric tests[2]. We conducted a 3: Group x 2: Colour x 2: Luminance mixed-design ANOVA with repeated measures over the last two factors. The results are displayed in panels B, C and F of Figure 2. The analysis revealed only one significant effect; Luminance, $F(1, 67) = 17.062$, $MSE = 0.1$, $p < .001$, $\eta_p^2 = .203$. Participants gave more long-short responses for intervals bounded by changes in luminance (M = 51.7%, 95% CI [49.1, 54.2]), than for intervals bounded by stimuli that did not change at all (M = 46.2%, 95% CI [43.7, 48.7]). There was a marginally significant main effect of Group, $F(2, 67) = 2.563$, $MSE = 3.2$, $p = .085$, $\eta_p^2 = .071$, with post-hoc comparisons using the Bonferroni correction finding that the Greek Short-Stay group made slightly more frequent long-short responses than the English group (p = .088), with no other contrasts yielding significant outcomes (ps > .4). Most crucially, however, was the absence of any evidence of a significant three-way interaction, $F(2, 67) = 1.524$, $MSE = 0.6$, $p = .225$, $\eta_p^2 = .044$. Crucially, Bayesian

---

[2] Given that the data for one cell in the Short-Stay group was not normally distributed, we conducted a parallel non-parametric test for this group alone. A Wilcoxon signed-rank test found no difference in the size of the increase in long-short responses for a change in green or a change in blue (p = .306), supporting the null finding in the parametric tests.

analyses found that the data were four times more likely under the null hypothesis that there is no three-way interaction, $BF_{10} = 0.257$, again showing meaningful support for the null by Dienes (2014) criterion. No other main effects or interactions reached significance: Colour, $F(1, 67) = 0.354$, $MSE = 0.8$, $p = .354$, $\eta_p^2 = .013$; Colour x Group, $F(2, 67) = 0.594$, $MSE = 0.9$, $p = .555$, $\eta_p^2 = .017$; Luminance x Group, $F(2, 67) = 0.095$, $MSE = 1.2$, $p = .909$, $\eta_p^2 = .003$; Colour x Magnitude, $F(2, 67) = 0.000$, $MSE = 0.6$, $p = .979$, $\eta_p^2 = .000$.

3.3 Are effects of the **ble/ghalazio** *boundary restricted to those who freely associate the stimuli with these terms?*

Another possibility is that only those Greek speakers who named the blue stimuli *ble* and *ghalazio* when no restrictions were applied to their colour naming choices at all will show an effect of crossing the *ble*/*ghalazio* boundary. For these participants, the stimuli clearly and uncontroversially belonged to these categories. We applied the same reasoning to this analysis as we did for the analysis incorporating length of stay, namely that a three-way interaction Group (English vs. Greek Strong Category vs. Greek Weak Category), Colour (Green vs. Blue), and Luminance (Same-Change vs. Same-Change) would support this hypothesis.

We conducted a 3: Group (English vs. Greek Weak Category vs. Greek Strong Category) x 2: Colour (Green vs. Blue) x 2: Luminance (Same-Change vs. Same-Change) mixed-design ANOVA with repeated measures over the last two factors again. The results are displayed across figures D, E and F in Figure 2. The analysis revealed only a significant main effect of Luminance, $F(1, 67) = 16.314$, $MSE = 1.2$, $p < .001$, $\eta_p^2 = .196$, indicating that participants gave more long-short responses for intervals bounded by stimuli that changed in luminance ($M = 51.6\%$, 95% CI [49.0, 54.2]), than stimuli that did not change (*M* = 46.1%, 95% CI [43.5, 48.7]). Again, no other main effects or interactions reached significance: Colour, $F(1, 67) = 0.749$, $MSE = 0.9$, $p = .390$, $\eta_p^2 = .011$; Group, $F(2, 67) = 1.611$, $MSE = 3.3$, $p = .207$, $\eta_p^2 = .046$; Colour x Group, $F(2, 67) = 0.292$, $MSE = 0.9$, $p = .748$, $\eta_p^2 = .009$; Luminance x Group, $F(2, 67) = 0.001$, $MSE = 1.2$, $p = .921$, $\eta_p^2 = .166$; Colour x Luminance, $F(2, 67) = 0.020$, $MSE = 0.6$, $p = .887$, $\eta_p^2 = .000$). Crucially, there was no three-way interaction, $F(2, 67) = 0.370$, $MSE = 0.6$, $p = .692$, $\eta_p^2 = .011$. Again, Bayesian analyses found that the data were approximately six times more likely under the null hypothesis that there was no three-way interaction, $BF_{10} = 0.158$, which once more exceeded Dienes' (2014) suggested criterion for a meaningful null result.

3.4. Additional analyses

The results of our analyses pointed strongly to there being no effect of crossing a colour category boundary on duration judgments. Since we were explicitly exploring null results, we conducted three final post-hoc tests in which we analysed the data from each of the native English speakers, the Greek Short-Stay group, and the Greek Strong-Category group separately using repeated-measures ANOVAs (Luminance x Colour). Note that we included these analyses only for completeness, since further tests were not mandated by the results of the previous ANOVAs. Here we report only the results relating to whether there was an interaction at the within-group level between Luminance and Colour, which would indicate that a change in the luminance of blues influenced duration judgments differently from a change in the luminance of greens. The full results can be found in the supplemental materials.

The analysis of the native English speakers ($n = 20$) revealed no interaction, $F(1, 19) = 0.417$, $MSE = 0.4$, $p = .526$, $\eta_p^2 = .021$. Bayesian analyses found that the data were approximately twice as likely under the null hypothesis that there was no Colour x Luminance interaction, $BF_{10} = 0.565$. The same analysis for the Greek Short-Stay group also found no Colour x Luminance interaction, $F(1, 24) = 6.107$, $MSE = 1.1$, $p = .29$, $\eta_p^2 = .047$, and again Bayesian analyses found that the data were two times more likely under the null that there is no interaction, $BF_{10} = 0.414$. Finally, the results of the same analysis with the Greek Strong Category group also found no Colour x Luminance interaction, $F(1, 18) = 0.134$, $MSE = 0.3$, $p = .719$, $\eta_p^2 = .007$, and Bayesian analyses found that the data were more than three times more likely under the null that there is no interaction, $BF_{10} = 0.305$, meeting Dienes' (2014) criterion even in this more restricted sample.

Next, we looked at the data from the blocks with numerals. Since we found strong support for the null that the *ble/ghalazio* boundary in fact had *no* special effect on duration perception, the analysis with numeral magnitude is largely included here for completeness. In order to provide the best comparison with our previous analyses, we conducted two mixed-design ANOVAs, in each case with Magnitude (One-Seven vs. Seven-One) as the within-subjects factor, but in one analysis the between-subjects factor Group was the same as in the previous analysis for length of stay (English vs. Greek Short-Stay vs. Greek Long-Stay), and in the other it was the same as the previous analysis for colour naming choices (English vs. Greek Strong Category vs. Greek Weak Category). The results of both analyses are displayed in Figure 3. In both analyses, there was a significant main effect of Magnitude: analysis by length of stay, $F(1, 67) = 6.880$, $MSE = 0.4$, $p = .011$, $\eta_p^2 = .093$; analysis by naming choices, $F(1, 67) = 6.840$, $MSE = 0.4$, $p = .011$, $\eta_p^2 = .093$, with more long-short responses when the first interval was bounded by numerals that differed by seven steps than by one step alone: analysis by length of stay ($M_{Diff} = 2.7\%$, 95% CI [0.1, 4.7]); analysis by naming choices (M_{Diff} = 2.7%, 95% CI [0.1, 4.8]). In both analyses there was no main effect of Group: analysis by length of stay, $F(2, 67) = 1.966$, $MSE = 2.3$, $p = .148$, $\eta_p^2 = .055$; analysis by naming choices, $F(2, 67) = 1.918$, $MSE = 2.3$, $p = .011$, $\eta_p^2 = .054$. In both cases there was also no evidence of any interaction: analysis by length of stay, $F(2, 67) = 0.257$, $MSE = 0.4$, $p = .774$, $\eta_p^2 = .008$; analysis by naming patterns, $F(2, 67) = 0.202$, $MSE = 0.4$, $p = .817$, $\eta_p^2 = .006$. In sum, the results from the blocks with numerals replicated the effect found by Alards-Tomalin and colleagues (2014), and patterned similarly to the results from the blocks that manipulated luminance.

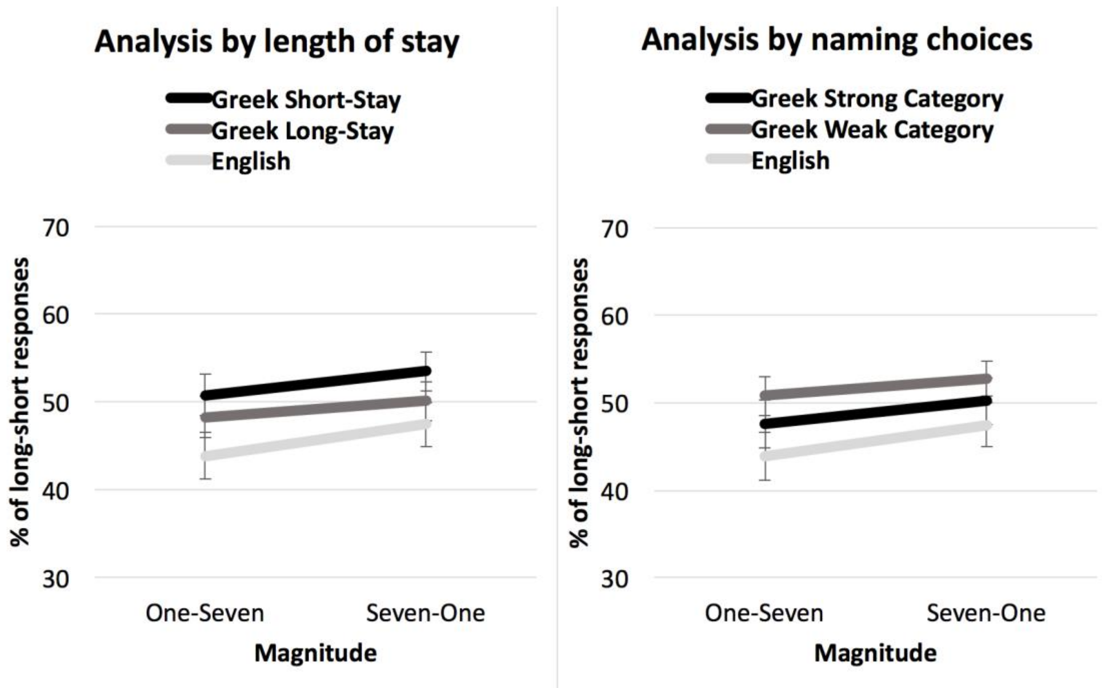

Figure 3. Results of analyses with numerals (with standard error bars).

4. Discussion

In the present experiment, we tested the limits of ATOM's notion of a shared and analogue metric for number, luminance and duration. We presented Greek and English speakers with a duration judgment task in which they had to indicate whether they perceived one interval to be either longer or shorter than the other. Consistent with previous research (Xuan et al., 2007), we found that changes in luminance influenced participants' judgments of duration, such that intervals bounded by stimuli that changed in luminance were perceived to last longer than intervals bounded by stimuli that did not change. Crucially, however, we found no evidence that crossing a category boundary had any effect on duration perception. Greek speakers showed no difference in performance when they perceived a change in blue, which for them consisted simultaneously of a change in colour category, relative to when they saw a change in green, or relative to changes in blue in native English speakers. Breaking down the Greek participants by their time spent in the UK, we also found no effect of acculturation to an English-language environment on performance on these stimuli. Similarly, and perhaps most tellingly, even those Greek speakers who preferentially associated the terms *ble* and *ghalazio* to the relevant stimuli in a completely free-choice naming task showed no special effect of crossing this category boundary. Overall, these results therefore support the hypothesis of a shared analogue representation of luminance and duration that is impervious to discrete changes contained within magnitude being perceived.

Although there are indeed many instances where language and cross-linguistic differences have been shown to influence cognitive processes, including in the temporal domain (see Athanasopoulos, Samuel, & Bylund, 2017, and Winter et al., 2015, for reviews), it would appear that there are constraints on this influence. The present results suggest that, to influence time perception, language may need to be specifically *temporal* language (e.g. related to 'time' vocabulary, see for example Casasanto, 2008), or alternatively may not be able to influence shared magnitude processing at all.

Overall, we believe our results are consistent with the suggestion in ATOM that there is a shared metric for luminance and duration. The present study also lends some support to the additional contention that this shared metric is common to all humans, and independent of experience (Walsh, 2003). Our results demonstrate that this analogue metric is not susceptible to interference from 'learned' linguistic categories. It could be that, as ATOM proposes, this shared metric is therefore available 'from birth'. In support of this possibility, there is evidence that very young infants and even neonates just a few hours old show evidence of shared magnitude processing (e.g. de Hevia, Izard, Coubart, Spelke & Steri, 2014; Lourenço & Longo, 2010; Srinivasan & Carey, 2010).

Our results would appear not to support the suggestion by Alards-Tomalin and colleagues (2014) that it is the integration of more proximal stimuli (e.g. 1, 2) that may underlie duration judgment effects in such tasks. Since the two blue stimuli should be more distinct for the native Greek speakers than for the native English speakers (cf. Thierry et al., 2009), and experientially we would expect that Greek speakers would have had many more occasions in their lifetime to distinguish between the two blue stimuli given that word choice represents a more fundamental decision in this population, we should have found that Greek speakers found it more difficult to integrate these categories, but we found no such evidence

Although there are other explanations of the effects of magnitude on duration judgments, it is important to note that these accounts were designed with the duration of stimuli themselves and *not* the intervals between them in mind. For example, the 'oddball' effect predicts that stimuli that are perceived as more distinct or less predictable in their context are more likely to have their duration of presentation overestimated (e.g. Schindel, Rowlands, and Arnold, 2011; Tse, Intriligator, Rivest, &

Cavanagh, 2004). The 'contrast' account posits that the magnitude of the contrast between the luminance of a stimulus and the luminance of its background may explain duration overestimation (e.g. Matthews, Stewart, & Wearden, 2011). The neural amplification hypothesis posits (among other predictions) that stimulus repetition leads to a decrease in neural response and hence a concomitant decrease in perceived duration (e.g. Eagleman, 2008; Pariyadath, & Eagleman, 2007, 2012). Although it is true that magnitude effects in AXB tasks rely on characteristics of the stimuli bounding the intervals, and hence such theories may indeed be relevant, at this stage we feel that there is no theoretical framework within these theories that makes any direct prediction about effects on empty intervals[3]. As a result, we do not argue that the results from the present study either support or contradict such hypotheses. Nevertheless, what is clear is that if such effects were impacted by the processes outlined above, they also did not show any evidence of interacting with CP.

Could our findings be explained by features of our sample of Greek speakers? It is certainly true that the Greek speakers in this experiment had a high level of English since they were studying in the UK at the time of testing (mean 4.28/5 self-reported proficiency). However, and as mentioned earlier, there is good evidence to suggest that our first language influences our processing even when we are engaged in

[3] It may be worth noting here that since the grey background on which the stimuli were displayed was closer in luminance to the darker colours than the lighter colours, it is unlikely that the contrast account could explain the effects of luminance on duration judgments here, since the effect was found regardless of whether the stimulus became more similar (i.e. darker) or more distinct (i.e. lighter) than the background (see supplemental materials for details of this analysis).

an explicitly second-language context; we cannot 'turn it off' (e.g. Marian & Spivey, 2003a, 2003b; Thierry & Wu, 2007; Wu & Thierry, 2010). Most relevant to the present study are the findings of Thierry et al.'s (2009) study that we described earlier, but also our own evidence that Greek speakers who had spent on average just four months in the UK performed no differently from those who had spent 40 months in the country. In sum, previous studies suggest that the Greek speakers' first language colour terms can permeate into tasks conducted in a second-language environment, and our own data together suggests that the degree of acculturation (or indeed non-acculturation) to an English-speaking environment does not modulate the effect of luminance change.

Secondly, our analyses of the specific naming choices of participants ruled out the possibility that we found no effect of the *ble/ghalazio* boundary because Greek speakers may not agree on which colours are best represented by the terms. All 50 native Greek speakers that were included in the present study named the two stimuli *ble* and *ghalazio* in a two-alternative forced-choice task. Nevertheless, the most stringent test is to look at the evidence from those Greek speakers who not only named the colours correctly in the forced-choice task, but also those who gave the same terms when no restrictions on their choices were applied. Crucially, there was no evidence that these participants performed differently when crossing this colour category boundary compared to when they did not, relative to both greens and to the other Greek speakers and native English speakers. It is worth pointing out also that conducting and using a free-choice naming test is a particularly stringent criterion that goes beyond what is typically found in other studies in the field. In sum, regardless of whether one prefers to rely on the full sample of Greek speakers or the subset who demonstrate the strongest possible adherence to the labels crucial to our hypothesis, there was no evidence of an effect of colour category boundaries on duration judgments.

Would it be reasonable to expect an effect of colour on duration judgments if we presented the stimuli to the right visual field, and hence favoured the categorical processing of colour in the language regions of the brain in the left hemisphere? This approach has proved fruitful in other studies of the categorical perception of colour (e.g. Roberson et al., 2008; Winawer et al., 2008). In these studies, the finding that participants are faster to locate an across-category target in the right visual field provided evidence that the idiosyncratic colour categories of language aided target location. However, these studies were concerned with the ability to *locate* a unique stimulus among competitors. It is clear from the results of the numeral stimuli in the present study, as well as previous studies (e.g. Alards-Tomalin et al., 2014) that participants have no problem perceiving the stimuli in AXB tasks, as evidenced by the fact that numerals can and do influence duration perception in the absence of lateralised presentation. Indeed, it is difficult to imagine a convincing argument to suggest that categorical differences can *only* be perceived if they are not presented centrally. For example, if instead of a light blue followed by two dark blues we had presented an image of a dog followed by two cats, it is clear that participants would recognise that the stimuli change categorically also in binocular view. Thus, our manipulation would not require the lateralisation of stimuli in order for any effect of the stimuli to occur.

4. Conclusion

According to ATOM, number, luminance, spatial and temporal codes share the same analogue magnitude metric, and hence behavioural experiments should uncover evidence of cross-talk between codes. We found evidence that this process is impervious to discrete changes in the stimuli from which the magnitude finds it source. As a result, the findings of the present study lend support to the suggestion that analogue luminance data is neither interrupted nor distorted by simultaneous and discrete (digital) shifts in the same input stream.

Acknowledgments

We are grateful to the department of psychology (in particular Jo Stephens and Devon Wybrow) and the ESSEXLab facility at the University of Essex, UK, for their assistance with this project.

Appendix

Colour measurements. Observed Yxy values for colour stimuli and grey background. The Y indicates the level of luminance of that stimulus.

	Y	x	y
Light Green	35.73	.27	.38
Dark Green	9.62	.26	.43
Light Blue	35.48	.24	.26
Dark Blue	9.15	.22	.23
Grey (background)	14.21	.31	.33

References

Abe, S. (1935). Experimental study of the correlation between time and space. *Tohoku Psychologica Folia*.

Alards-Tomalin, D., Leboe-McGowan, L. C., & Mondor, T. A. (2013). Examining auditory kappa effects through manipulating intensity differences between sequential tones. *Psychological research*, *77*(4), 480–491.

Alards-Tomalin, D., Leboe-McGowan, J. P., Shaw, J. D., & Leboe-McGowan, L. C. (2014). The effects of numerical magnitude, size, and color saturation on perceived interval duration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 555–566.

Alards-Tomalin, D., Walker, A. C., Kravetz, A., & Leboe-McGowan, L. C. (2015). Numerical Context and Time Perception: Contrast Effects and the Perceived Duration of Numbers. *Perception*, *45*(1-2), 222–245.

Androulaki, A., Gômez-Pestaña, N., Mitsakis, C., Jover, J. L., Coventry, K., & Davies, I. (2006). Basic colour terms in modern Greek: Twelve terms including two blues. *Journal of Greek Linguistics*, *7*(1), 3–47.

Athanasopoulos, P. (2009). Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition*, *12*(01), 83–95.

Athanasopoulos, P., Dering, B., Wiggett, A., Kuipers, J. R., & Thierry, G. (2010). Perceptual shift in bilingualism: Brain potentials reveal plasticity in pre-attentive colour perception. *Cognition*, *116*(3), 437–443.

Athanasopoulos P., Samuel S., Bylund E. (in press). The psychological reality of spatio-temporal metaphors. In *Studies in Figurative Thought and Language*, A. Athanasadiou, P. (Ed). Amsterdam: John Benjamins.

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1525), 1831–1840.

Casasanto, D. (2008). Who's afraid of the big bad Whorf? Crosslinguistic differences in temporal language and thought. *Language learning*, *58*(s1), 63–79.

Cohen, J., Hansel, C. E. M., & Sylvester, J. D. (1954). Interdependence of temporal and auditory judgments. *Nature*, *174*, 642-644.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. Frontiers in Psychology, 5, 781.

Dormal, V., Seron, X., & Pesenti, M. (2006). Numerosity-duration interference: A Stroop experiment. *Acta psychologica*, *121*(2), 109–124.

Droit-Volet, S., & Gil, S. (2009). The time–emotion paradox. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1525), 1943–1953.

Eagleman, D. M. (2008). Human time perception and its illusions. *Current opinion in neurobiology*, *18*(2), 131–136.

de Hevia, M. D., Izard, V., Coubart, A., Spelke, E. S., & Streri, A. (2014). Representations of space, time, and number in neonates. *Proceedings of the National Academy of Sciences*, *111*(13), 4809–4813.

Lourenco, S. F., & Longo, M. R. (2010). General magnitude representation in human infants. *Psychological Science, 21*(6), 873–881.

Marian, V., & Spivey, M. (2003a). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition*, *6*(2), 97–115.

Marian, V., & Spivey, M. (2003b). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics*, *24*(2), 173–193.

Matthews, W. J., Stewart, N., & Wearden, J. H. (2011). Stimulus intensity and the perception of duration. *Journal of Experimental Psychology: Human perception and performance*, *37*(1), 303.

Oliveri, M., Vicario, C. M., Salerno, S., Koch, G., Turriziani, P., Mangano, R., Chillemi, G., & Caltagirone, C. (2008). Perceiving numbers alters time perception. *Neuroscience letters*, *438*(3), 308–311.

Ornstein, R. (1969). *On the experience of time*. Hammondsworth: Penguin.

Pariyadath, V., & Eagleman, D. (2007). The effect of predictability on subjective duration. *PloS one*, *2*(11), e1264.

Pariyadath, V., & Eagleman, D. M. (2012). Subjective duration distortions mirror neural repetition suppression. *PLoS One*, *7*(12), e49362.

Pöppel, E. (2009). Pre-semantically defined temporal windows for cognitive processing. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1525), 1887–1896.

Poynter, W. D., & Homa, D. (1983). Duration judgment and the experience of change. *Perception & Psychophysics*, *33*(6), 548–560.

Roberson, D., Davidoff, J., Davies, I. R., & Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive psychology*, *50*(4), 378–411.

Roberson, D., Pak, H., & Hanley, J. R. (2008). Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition*, *107*(2), 752–762.

Sarrazin, J. C., Giraudo, M. D., Pailhous, J., & Bootsma, R. J. (2004). Dynamics of balancing space and time in memory: tau and kappa effects revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(3), 411–430.

Schindel, R., Rowlands, J., & Arnold, D. H. (2011). The oddball effect: Perceived duration and predictive coding. *Journal of vision*, *11*(2), 17.

Srinivasan, M., & Carey, S. (2010). The long and the short of it: on the nature and origin of functional overlap between representations of space and time. *Cognition*, *116*(2), 217–241.

Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J. R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences*, *106*(11), 4567–4570.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, *104*(30), 12530–12535.

Tse, P. U., Intriligator, J., Rivest, J., & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception & psychophysics*, *66*(7), 1171–1189.

Vicario, C. M., Pecoraro, P., Turriziani, P., Koch, G., Caltagirone, C., & Oliveri, M. (2008). Relativistic compression and expansion of experiential time in the left and right space. *PLoS One*, *3*(3), e1716.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in cognitive sciences*, *7*(11), 483–488.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, *104*(19), 7780-7785.

Winter, B., Marghetis, T., & Matlock, T. (2015). Of magnitudes and metaphors: Explaining cognitive interactions between space, time, and number. *Cortex*, *64*, 209–224.

Wu, Y. J., & Thierry, G. (2010). Chinese–English bilinguals reading English hear Chinese. *Journal of Neuroscience*, *30*(22), 7646–7651.

Xuan, B., Zhang, D., He, S., & Chen, X. (2007). Larger stimuli are judged to last longer. *Journal of vision*, *7*(10), 2.