Introducing a control condition in the classic oddball paradigm: Oddballs are overestimated in duration not only because of their oddness

Teresa Birngruber · Hannes Schröter · Rolf Ulrich

Abstract Three experiments assessed whether participants perceive infrequent stimuli (oddballs) to last longer than frequent ones (standards). The classical oddball paradigm was modified so that participants judged the duration of a target stimulus which could either be a standard or an oddball. The target was always the fifth stimulus in the stream and all stimuli were presented at predefined spatial positions. These modifications enabled a direct comparison of duration judgments for oddballs and standards. In Experiments 1 and 2 not only the duration of oddballs but also the duration of standards was overestimated by virtually the same amount. In other words, the overestimation of oddballs was not due to their oddness but reflected a different temporal dilation such as the negative time-order error. In Experiment 3, all stimuli were presented at the same spatial position. Again, both oddballs and standards were overestimated, however, oddballs more so. The present results highlight the importance of comparing the judged duration of oddballs and standards when evaluating the size of the genuine oddball effect. Furthermore, the findings suggest that the occurrence of temporal oddball effects can depend on spatial features of stimulus presentation.

Keywords Time perception · Oddball effect · Stimulus frequency

Previous research has indicated that people tend to overestimate the duration of deviant stimuli appearing in a series of repeated standard stimuli (Birngruber et al. 2014a; Chen and Yeh 2009; Kim and McAuley 2013; New and Scholl 2009; Pariyadath and Eagleman 2007; Schindel et al. 2011;

T. Birngruber (✉) · H. Schröter · R. Ulrich
University of Tübingen, Tübingen, Germany
e-mail: teresa.birngruber@uni-tuebingen.de

Tse et al. 2004). Tse et al. first introduced a paradigm that revealed this oddball effect (OE). In this paradigm, participants were presented a stream of standard stimuli, for example small black disks, in the center of a computer screen. All standard stimuli were presented for a constant standard duration. At random positions within this stream, single oddball stimuli, for example expanding black disks, were presented for varying comparison durations. The participants' task was to judge whether the oddballs were shorter or longer in duration than the standard stimuli. Tse et al. observed overestimations of the oddballs' duration in the order of 35 %. This basic effect has not only been observed for expanding disks but also for stationary oddballs which only deviated from the standards in shape or color (Birngruber et al. 2014a; Chen and Yeh 2009; Schindel et al. 2011).

However, follow-up studies have usually reported distinctly smaller OEs than those reported by Tse et al. (2004). In their original study, Tse et al. employed comparison durations that were mostly shorter than the standard duration. Such an asymmetric arrangement of comparison durations can result in the so-called asymmetry effect (Guilford 1954) and thus artificially inflate the size of the OE. Accordingly, Seifried and Ulrich (2010) observed smaller OEs of approximately 5 % when comparison durations were arranged symmetrically around the standard duration.

An OE has also been observed in a modified paradigm, in which two successive stimuli were presented in each trial (Ulrich et al. 2006). In this paradigm, a visual standard stimulus of constant duration always preceded a visual target stimulus of varying duration. There were two types of comparison stimuli. A frequent comparison appeared in 70 % of the trials, whereas nine different infrequent comparisons appeared in the remaining 30 % of trials. Consequently, the presentation probability of a single infrequent

comparison was only 3.3 %. Participants judged whether the second stimulus duration appeared longer or shorter than the first stimulus duration. The authors observed small (approximately 10-25 ms) but significant overestimations of infrequent comparisons as compared to frequent ones for stimulus durations of 400 and 800 ms. While this modified paradigm allowed the effect of stimulus probability on judged duration to be isolated, the underlying mechanism of this frequency effect might be quite different from that of the classic stream-based OE. Further evidence for the idea that the stream might be processed differently comes from the observation that the size of the OE normally increases with the position of the oddball in the stream (Birngruber et al. 2014a; Kim and McAuley 2013; Pariyadath and Eagleman 2012).

Within the classic stream-based paradigm, however, the commonly used quantification of the OE might be problematic. Usually a psychometric function is determined for each participant from which the point of subjective equality (PSE) for oddball stimuli is estimated. The PSE denotes the comparison duration at which oddballs and standards are judged as being equally long. The OE is then quantified as the difference between the PSE and the objective standard duration. Thus, an OE measured this way must not necessarily result from an overestimation of oddballs but could also reflect an underestimation of standards (see also, Matthews, 2011). In addition, this measure could be contaminated by a negative time-order error, that is, the tendency for participants to overestimate the second stimulus when two successive stimuli must be compared (Allan 1977; Jamieson and Petrusic 1975; Woodworth and Schlossberg 1965, p. 226). Because of these reasons, it is desirable to estimate a separate PSE for judgments about standard stimuli in addition to the PSE for oddballs. Accordingly, the PSE for oddballs should be smaller than the PSE for standards, if participants perceive oddballs to be longer than standards.

Van Wassenhove et al. (2008) introduced a modified stream-based paradigm in which this direct comparison of PSEs for oddballs and PSEs for standards was possible. The authors, however, focused on distortions of duration perception within and across sensory modalities rather than on the OE per se. They presented a stream of five stimuli in each trial, and participants always judged the duration of the fourth stimulus. This target stimulus could either be another standard stimulus in the control block or a deviant stimulus in the experimental block. Consequently it was possible to estimate separate PSEs for standards and for deviants. Van Wassenhove et al. observed differences between the two PSEs, indicating that looming oddballs were overestimated relative to steady standards whereas steady oddballs were underestimated relative to looming standards. Thus, these authors observed results which contrast with the findings of

Tse et al. (2004) who reported overestimations of oddballs in both cases. A potential complication of this paradigm, however, is that participants had to count along in order to unmistakeably identify the target, especially in the control condition where only the target's duration but not its appearance was deviant from the standard stimuli. This might have resulted in different processing strategies for the experimental and the control blocks. Additionally, counting might generally influence duration estimates (Grondin et al. 2004; Rattat and Droit-Volet 2012).

Below, we introduce a paradigm that aims to address the aforementioned critiques of the stream-based oddball paradigm. To this end, we modified the procedure in a way that allowed for a direct comparison between the PSE for standards and the PSE for oddballs presented at the same position within the stream. In order to distinguish this scoring from the original measurement of the oddball effect, we will refer to this difference of PSEs as the *subjective oddball effect* (SOE). To render counting unnecessary, stimuli were presented in a spatially predictable manner, namely clockwise around an orientation point. This enabled the identification of the target stimulus based only on its spatial location (i.e., the 12 o'clock position).

As well as addressing these basic methodological issues, the following experiments aimed to provide insight into the conditions that might modulate the size of the oddball effect. First, oddballs, by definition, deviate from the repeated standards in a certain feature, for example color or shape. This may already be sufficient to trigger an SOE. Second, a low overall occurrence probability of oddballs might be especially crucial to elicit an overestimation of oddballs compared to standards. Finally, the conditional probability for an oddball to appear at the target position might in fact be the factor that contributes to the size of the SOE. Although these factors can not be manipulated orthogonally, it is nevertheless possible to evaluate the importance of these factors across experiments.

Experiment 1

Experiment 1 was a modification of the classic stream-based oddball paradigm. In each trial, five visual stimuli were presented in succession, four identical standard stimuli of constant duration followed by a target stimulus of varying duration (see Fig. 1A). This target could either be another standard or an oddball. The participants' task was to judge whether the target was shorter or longer than the preceding standards. In order to clearly identify the to-be-judged stimulus (without the need to count), stimuli were arranged in a spatially predictable, circular order and the target was always presented at the 12 o'clock position. Additionally,
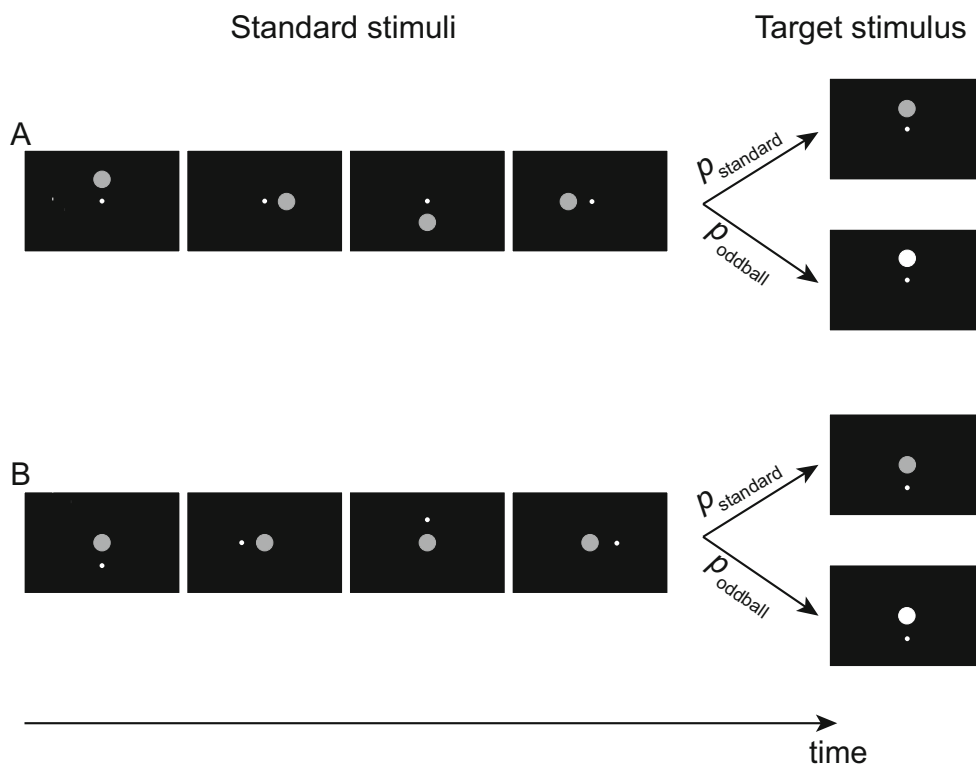
Standard stimuli   Target stimulus

Fig. 1 Schematic illustration of the trial procedure in (**A**) Experiments 1 and 2 and (**B**) Experiment 3. In Experiments 1 and 2, the stimuli were presented in a clockwise manner around the centrally presented orientation point. In Experiment 3, the stimuli were presented centrally with the orientation point moving around the stimuli. The standard stimuli (depicted in gray) were either blue or red (for half of the participants, respectively) in all experiments. All standard stimuli were presented for the fixed standard duration of 800 ms. The target stimulus could either be another standard (depicted in gray) or an oddball (depicted in white). Target stimuli were presented for one of nine comparison durations. In Experiment 1 the probability for a standard target ($p_{standard}$) and for an oddball target ($p_{oddball}$) was 50 % each. In the single condition of Experiment 1, oddball targets were always red (if standards were blue) or blue (if standards were red). In the multi condition of Experiment 1, oddball targets were cyan, green, yellow, or magenta; each oddball color was presented with a probability of 12.5 %. In Experiments 2 and 3, $p_{standard}$ was 70 % and $p_{oddball}$ was 30 %. These 30 % were equally distributed across nine oddball stimuli which were created by combining the shapes square, triangle, and diamond with the colors green, yellow, and magenta. The probability for each individual oddball stimulus was thus 3.3 %

the frequency of the oddball color was manipulated between blocks to examine whether the SOE is modulated by stimulus probability. In the single condition, the target stimulus was a standard in half of the trials and one specific color oddball in the remaining half of the trials. Thus, the occurrence of a standard or an oddball at the target position was equally likely. In the multi condition, the target was again a standard in 50 % of trials. In each remaining trial, one of four different color oddballs was presented with equal probability. Thus, the effective probability of a specific oddball at the target position was only 12.5 % in this condition. If the mere deviation of the oddball from the standards is crucial for the occurrence of the SOE, similar effect sizes should be observed in both conditions. However, if the probability of a specific stimulus modulates the size of the SOE, one should expect a larger SOE in the multi than in the single condition.

The PSE was supplemented by two additional measurements. First, a measurement of discrimination sensitivity

was employed. As Ulrich et al. (2006) observed better temporal discrimination for infrequent than for frequent stimuli, here we investigated the robustness of this effect, which is of theoretical importance in understanding the mechanisms underlying the oddball effect (see Ulrich et al., 2006, for a discussion of this issue). Second, an implicit measurement of judgment certainty was derived from response times. In particular, response times were measured as a function of target duration. Research on two-choice discrimination performance (e.g., Henmon 1910; Smith and Vickers 1988; Vickers and Packer 1982) suggests that response times should be shortest for very short or very long target durations, because participants are most certain about whether to judge the target as being shorter or longer than the standards, respectively. For target durations closer to the standard duration, response times are expected to increase. Therefore the peak of this function should capture the objective target duration at which participants were most uncertain about their judgment, i.e. the target duration which seemed most

similar to the standards. This measure should thus accord with PSE results. We assume that even though the instructions only emphasized precision and not speed, response times might still contain valuable information about the ease of making judgments.

Method

Participants Forty members of the University of Tübingen (30 female, 18 - 43 years, $M = 23.4$ years) participated in the experiment. Thirty-six of them reported being right-handed; all had normal or corrected-to-normal visual acuity and normal color vision. The experimental session lasted approximately 75 minutes and participants either received partial course credit or were paid € 10. Seven additional students participated, but their data had to be excluded from analyses due to flat psychometric functions (2 participants), due to more than five missed catch trials in total (4 participants), or due to misunderstanding the task (1 participant).

Apparatus, stimuli, and durations The experiment was programmed in MATLAB® using the Psychophysics Toolbox 3 (Brainard 1997; Pelli 1997). A PC controlled stimulus presentation and recorded the participants' responses. The computer screen (standard VGA screen) had a resolution of 1024×768 pixels and a refresh rate of 150 Hz. The "X" and "M" keys and the space bar of a standard German keyboard served as response keys.

Filled circles (diameter 1.1^o of visual angle) of six different colors presented against a black background served as stimuli. The six colors were blue, red, cyan, green, yellow, and magenta. Their luminance was about $10\ cd/m^2$ (black $0.4\ cd/m^2$). For half of the participants, standards were blue and frequent oddballs were red; for the other half, standards were red and frequent oddballs were blue. For all participants, the infrequent oddballs were cyan, green, yellow, and magenta. Catch trial stimuli were presented in the standard color but with lower luminance (70 % of the standard luminance). A small white dot presented in the middle of the screen served as orientation point, and the stimuli were presented on an imagined clock face around this point. The distance between the orientation point and the stimuli was fixed to 3 cm.

The standard stimuli were always presented for a constant duration of 800 ms. Nine comparison durations were arranged symmetrically around that standard duration (as it has been suggested by Seifried and Ulrich 2010): 480, 560, 640, 720, 800, 880, 960, 1,040, and 1,120 ms. All interstimulus intervals and intertrial intervals were chosen randomly for each presentation from a pool of ten durations, ranging

from 560 to 740 ms in steps of 20 ms. All durations were chosen to be integer numbers of screen refreshes.

Procedure The experiment took place in a dimly lit and sound-attenuated room. Participants received written and verbal instructions before the experiment. The experiment comprised a total of 340 trials. The experimental session was divided into two parts, each composed of one practice block and eight experimental blocks. The practice block included eight experimental trials and two catch trials. Each experimental block included 18 experimental trials and two catch trials. Practice trials and catch trials were excluded from data analysis; thus, 288 trials were included in the final data analysis. Breaks followed each of the blocks and participants could terminate these breaks by pressing the space bar. However, a fixed break of five minutes separated the two parts of the experiment.

Each trial started with the presentation of the orientation point in the center of the screen. While the orientation point stayed in place, four standard stimuli were presented one after each other in a clockwise manner. Thus, the first stimulus appeared at the 12 o'clock position, the second stimulus at 3 o'clock, the third stimulus at 6 o'clock, and the fourth stimulus at the 9 o'clock position. All standards were separated by randomly chosen inter-stimulus intervals. After the fourth standard stimulus and another inter-stimulus interval, the target stimulus was presented at the 12 o'clock position. The target stimulus could either be another standard stimulus (in 50 % of trials) or an oddball stimulus (in the other 50 % of trials). This target was presented for one of the nine comparison durations. The orientation point then disappeared together with the target stimulus and the screen stayed black until a response was given. Participants were instructed to judge whether the target stimulus was shorter or longer than the four standard stimuli, irrespective of target color. After pressing the corresponding key, the next trial started with the presentation of the orientation point after a randomly-chosen inter-trial interval.

The two parts of the experiment differed concerning the employed oddballs. In the single condition, only one oddball color was used, namely red (when standards where blue) or blue (when standards were red). In the multi condition, four infrequent oddball colors were employed. In both conditions, target stimuli were standards in 50 % of the trials and oddballs in the remaining 50 % of trials. Consequently, in the multi condition, each of the four oddball colors appeared in 12.5 % of trials. Comparison durations were balanced across target stimuli (in both conditions) and across oddball color (in the multi condition). In practice blocks, only four comparison durations

were used: 480, 640, 960, and 1,120 ms with each target stimulus.

Catch trials were included to ensure that participants followed the standard stimuli with their gaze. In each catch trial, either the second, third, or forth standard stimulus was presented with a slightly lower luminance. The position of these catch stimuli was chosen randomly for each catch trial. Participants were instructed to press the space bar as soon as they detected such a catch stimulus. If participants responded within 3,000 ms, the screen turned black immediately after the key press and the next trial started after an inter-trial interval. If participants failed to respond within this time window, a short message appeared on the screen reminding the participants to press the space bar as quickly as possible whenever they detected a stimulus with lower luminance.

Design and data analysis The experiment had a 2×2 factorial structure, resulting from the orthogonal combination of the within-subject factors Target Stimulus (standard vs. oddball) and Oddness Condition (single vs. multi). The colors of standards and oddballs, the judgment-to-key assignment, and the order of oddness conditions were counterbalanced across participants.

The PSE, indicating the specific comparison duration that was perceived as being equal to the standard duration, was the main dependent variable. It was determined individually and for each condition with the non-parametric Spearman-Kärber method (Lapid et al. 2008; Miller and Ulrich 2001). This technique enables one to determine a measure of location of the psychometric function without any specific assumptions about the shape of this function. This location parameter serves as the PSE estimate. The SOE was defined as the difference between the PSE for oddballs and the PSE for standards, $SOE = PSE_{oddballs} - PSE_{standards}$. As a measure of discrimination sensitivity, individual dispersion parameters from the Spearman-Kärber method served as estimates for the Difference limen (DL).

Furthermore, we analyzed response times (RT). To this end all RTs that were longer than 4,000 ms were excluded because these RTs seemed to be outliers (71 trials of the total of 11,520 trials were excluded which is less than 1 %). Besides the mean RTs for standard and oddball targets, we additionally investigated RTs as a function of comparison durations. We consider this an implicit measurement of judgment certainty since it is known that RT increases the more difficult the stimuli are to discriminate (Birren and Botwinick 1955; Smith and Vickers 1988). The mean RTs as a function of comparison durations formed an inverted U-shaped function illustrating that participants were slower to

judge targets that were similar to the standard duration than very short and very long targets. We determined the first moment of this comparison duration-RT function, using the Waveform Moment Analysis (Cacioppo and Dorfman 1987; Ulrich et al. 1995). This moment represents the point on the x-axis at which the mean of a bell-shaped function is located (see Birngruber et al. 2014a and Dyjas and Ulrich 2014, for other applications).[1] We used this procedure to infer the physical duration at which the participants were maximally uncertain about their judgments. We will refer to this comparison duration as the *Point of maximal uncertainty* (PMU). If the PSEs for standard and oddball targets differ, the two PMUs should also differ. We determined PMUs for each participant and each condition separately.

Results

Figure 2 displays the relative frequencies of "longer"-judgments for the nine comparison durations and the two oddness conditions (Panel A and B), averaged across all participants.

Point of subjective equality A two-factorial repeated measures ANOVA with the factors target stimulus and oddness condition was conducted on PSE. An SOE would be reflected by a significant main effect of target stimulus with smaller PSEs for oddball targets than for standard targets. No such main effect of target stimulus was present, $F(1, 39) = 1.94$, $MSE = 1,319$, $p = .171$, $\eta_p^2 = .05$, indicating that mean PSEs for standards and oddballs did not differ significantly (standard targets: 776 ms, oddball targets: 768 ms).[2] The main effect of oddness condition was not significant either, $F < 1$, signalling that PSEs were virtually identical in the two conditions (single: 771 ms, multi: 772 ms). The ANOVA revealed no interaction of the two factors, $F < 1$. Panel C of Fig. 2 depicts mean PSEs as a function of the two factors.

In order to examine whether the observed mean PSEs for oddball and standard targets differed from the objective standard duration, we conducted separate one-sample

---

[1]In brief, let m_i be the mean RT at target duration d_i, $i = 1, ..., 9$. In a first step, these means are scaled as $m_i^* = \frac{m_i}{\sum_{i=1}^9 m_i}$, $i = 1, ..., n$. In a second step, these scaled values are used to compute $M = \sum_{i=1}^9 d_i \cdot m_i^*$, that is, the location of the observed comparison duration-RT function.

[2]A power analysis revealed that with an alpha level of .05, a SOE of 25 ms, and a sampling variability of 46 ms (resembling the standard deviation of the SOEs in this experiment), the statistical power of obtaining a significant result is larger than 96 %.
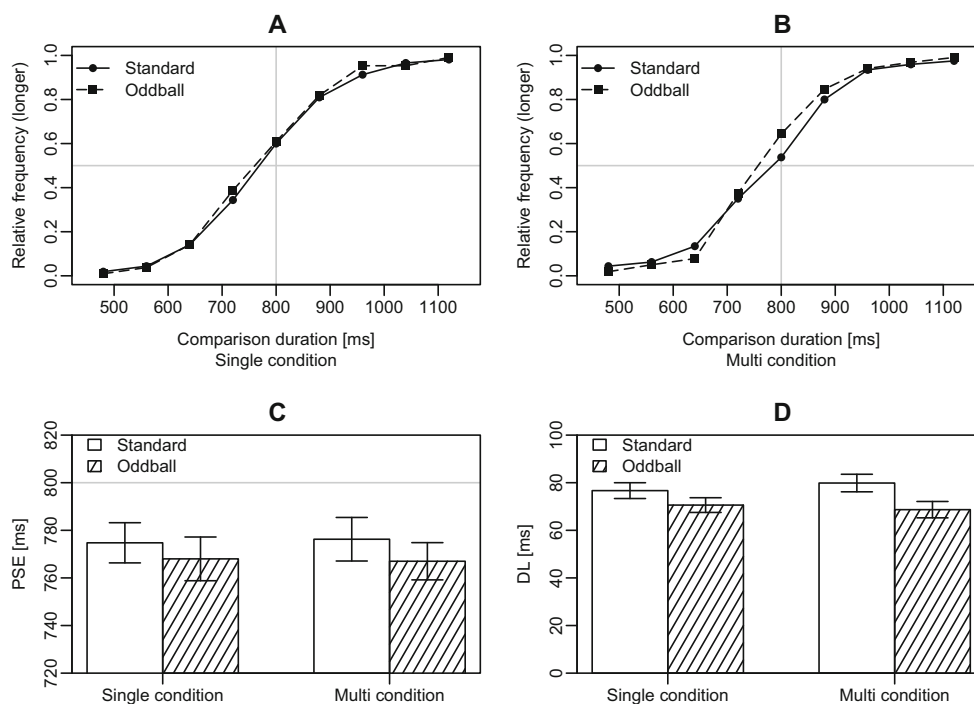
Fig. 2 Relative frequencies for longer responses as a function of comparison duration and target stimulus in the single **A** and multi **B** conditions of Experiment 1. Circles (squares) indicate relative frequencies in the case of a standard (oddball) target, averaged across participants. The horizontal gray lines indicate the probability for a longer judgment of 0.5. The vertical gray lines indicate the objective standard duration. **C** Point of subjective equality (PSE) as a function of oddness condition and target stimulus. **D** Difference limen (DL) for standard and oddball targets in the single and multi condition. DLs were calculated individually and are depicted averaged across participants. All error bars indicate ±1 standard error

t-tests. These tests represent the classic definition of the OE (when only oddball stimuli are judged). The mean PSE for oddball targets was statistically smaller than the objective standard duration of 800 ms, $t(39) = 4.24$, $p < .001$. The same was true for standard targets, $t(39) = 3.21$, $p = .003$. Thus, the durations of all target stimuli were significantly overestimated.

Difference limen Panel D of Fig. 2 depicts mean DLs as a function of oddness condition and target stimulus. The corresponding repeated measures ANOVA revealed a significant main effect of target stimulus, $F(1, 39) = 17.29$, $MSE = 172$, $p < .001$, $\eta_p^2 = .31$, indicating larger DLs for standard targets (78 ms) than for oddball targets (70 ms). This finding is consistent with results reported by Ulrich et al. (2006). An absent main effect of oddness condition showed that discrimination sensitivity did not differ between the two oddball probabilities, $F < 1$. The interaction between the two factors was not significant, $F(1, 39) = 1.57$, $MSE = 167$, $p = .218$, $\eta_p^2 = .04$.

Response time Panel A and B of Fig. 3 display the mean RTs for standard and oddball targets as a function of comparison duration in the two oddness conditions. As we expected, mean RTs followed an inverted U-shaped function

with RTs being longest at medium comparison durations. A repeated measures ANOVA revealed a significant main effect of target stimulus on RT, $F(1, 39) = 32.43$, $MSE = 4,781$, $p < .001$, indicating generally faster responses to oddballs (725 ms) than to standards (787 ms). Neither a significant main effect of oddness condition nor an interaction of the two factors was found (both $Fs < 1$).

Point of maximal uncertainty Individual PMUs were determined as described above and are shown in Panel C of Fig. 3. PMUs were very similar in all four conditions (standard, single: 778 ms; oddball, single: 775 ms; standard, multi: 777 ms; oddball, multi: 776 ms); the repeated measures ANOVA revealed no significant effects (all $Fs < 1$). A one-sample t-test showed that the grand mean PMU (776 ms) was significantly smaller than the objective standard duration of 800 ms, $t(39) = 10.12$, $p < .001$, which is in line with the PSE results. This indicates that participants were most uncertain about their duration judgment at target durations around 776 ms, irrespective of condition.

Catch Trials Catch stimuli were successfully detected in 99.4 % of all catch trials. Therefore, we can be confident
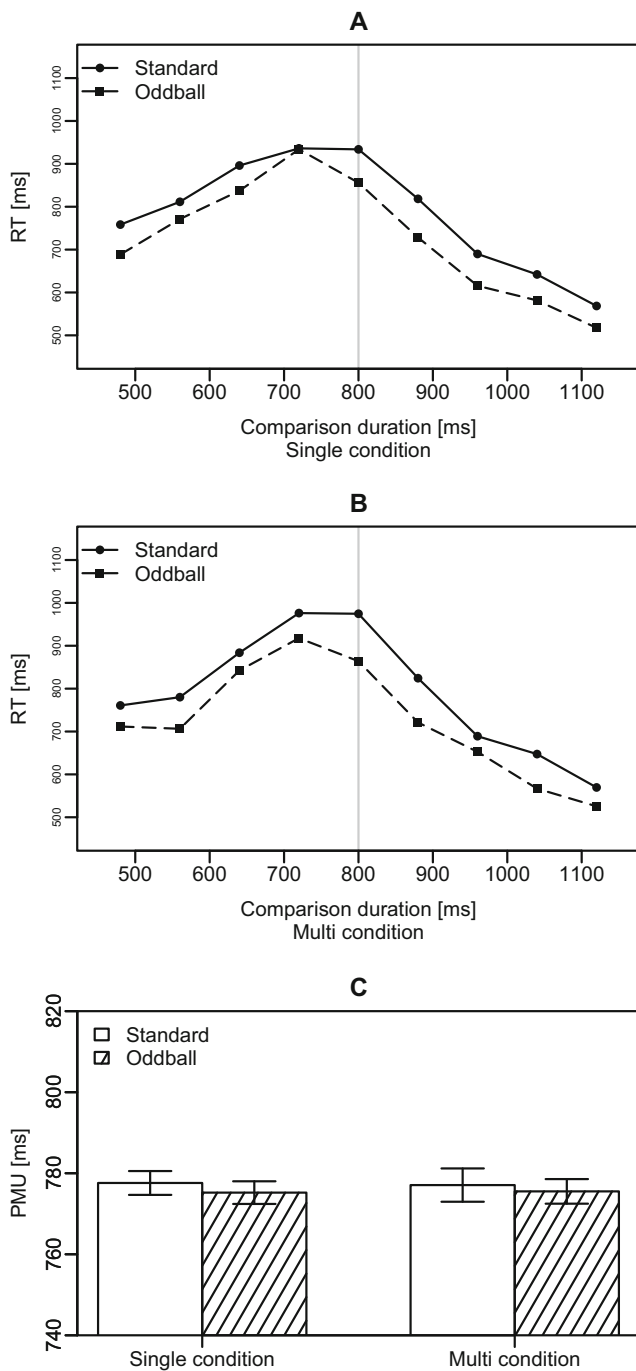
Fig. 3 Mean response times (RT) for standard and oddball targets in the single (**A**) and multi (**B**) condition of Experiment 1, as a function of comparison duration. The vertical gray lines indicate the objective standard duration of 800 ms. **C** Point of maximal uncertainty (PMU) as a function of oddness condition and target stimulus. These values indicate the target duration at which participants were most uncertain about their judgments. Error bars indicate ±1 standard error

that participants generally attended to the stimuli. The mean RT in catch trials, measured from the start of the catch stimulus to the key press was 954 ms.

Discussion

Experiment 1 revealed no difference between the subjective duration of oddball and standard targets. Therefore, no SOE was present in the data. Nevertheless, a general overestimation of all targets was observed. These findings were additionally supported by PMU data which revealed that responses were slowest at target durations around the PSE but this new measurement of judgment uncertainty also did not differ for standards and oddballs. This pattern of results is inconsistent with the notion that merely the oddballs' visual deviation from the standards is responsible for their temporal overestimation. Further, it shows that a low overall probability of oddballs is not sufficient to induce an oddball effect. If the probability of an oddball in the target position is assumed to be crucial for the occurrence of a SOE, an effect should emerge in the multi but not in the single condition of the experiment. However, no indication of an SOE was present in either oddness condition.

Although target stimuli did not differ concerning their PSEs, discrimination sensitivity was higher for oddball targets than for standard targets. This result replicated the result reported by Ulrich et al. (2006) who attributed this sensitivity effect to an increased arousal-level induced by infrequent stimuli. This arousal account may also explain why mean RTs are shorter for oddballs than for standards. Nevertheless, the arousal account would also predict a SOE.

One possible explanation for the absence of the SOE might be that participants indeed took the probability of an oddball appearing at the target position into account but simply grouped the four colored oddballs in the multi condition to one category, i.e., "not the standard color". The probability for this category to occur in the target position would still have been 50 %. To investigate this possibility, we conducted a second experiment, in which we reduced the overall probability for an oddball in the target position and additionally increased the difference between specific oddballs in order to make it harder for the participants to group them together as one category.

Experiment 2

Oddness was manipulated in two ways in Experiment 2. First, target stimuli were oddballs in only 30 % of trials, while 70 % of trials comprised standard stimuli as targets. Second, combinations of three oddball colors and three oddball shapes served as oddball stimuli. Consequently, the frequency of each specific oddball stimulus was further reduced in comparison to the multi condition of Experiment 1. These oddball manipulations were adapted from the experiments of Ulrich et al. (2006), in which a reliable

difference between the PSEs for expected and unexpected comparisons was observed.

Method

Participants A fresh sample of forty members of the University of Tübingen (30 female, 18 - 58 years, $M = 23.9$ years) participated in the experiment. Thirty-three of them were right-handed; all had normal or corrected-to-normal visual acuity and normal color vision. The experimental session lasted approximately 90 minutes and participants either received partial course credit or were paid € 12. Three additional people participated, but their data had to be excluded from analyses due to flat psychometric functions.

Apparatus, stimuli, and durations The same equipment was used as in Experiment 1. Besides the standard stimuli from Experiment 1 (the red or blue circles), three additional geometrical shapes served as oddballs: an isosceles triangle, a square, and a rhombus. All shapes were matched to have approximately the same area as the circles. The standard colors were again blue and red and the oddball colors were now green, yellow, and magenta. The combination of the three oddball shapes and the three oddball colors resulted in nine oddball stimuli.

Procedure The experimental session started with a practice block of ten trials (including two catch trials) which was followed by four experimental blocks of 100 trials each (including 90 experimental trials and ten catch trials). Experiment 2 differed from Experiment 1 mainly concerning the presentation probabilities of the different target stimuli (see Fig. 1). Standard targets were presented in 70 % of the experimental trials (63 trials per experimental block) whereas oddball targets were presented in 30 % of the experimental trials (27 trials per experimental block). Each oddball shape appeared nine times per experimental block, the oddball colors were chosen randomly per oddball trial. Oddball colors were randomized because complete balancing with all nine oddball stimuli would have resulted in too many trials (300) per experimental block. The average occurrence probability for each of the nine oddball stimuli was thus $0.3 \cdot 0.111 = 0.033$. The order and arrangement of stimuli was otherwise identical to Experiment 1. In total, 360 experimental trials entered the final data analysis (practice trials and catch trials were excluded from data analysis). Short breaks were integrated after every 20 trials. A five minute break occurred in the middle of the experiment.

Design and data analysis This experiment was a one-factorial design with the within-subject factor Target Stimulus (standard target vs. oddball target). PSEs, SOEs, DLs, and PMUs were calculated analogously to Experiment 1.
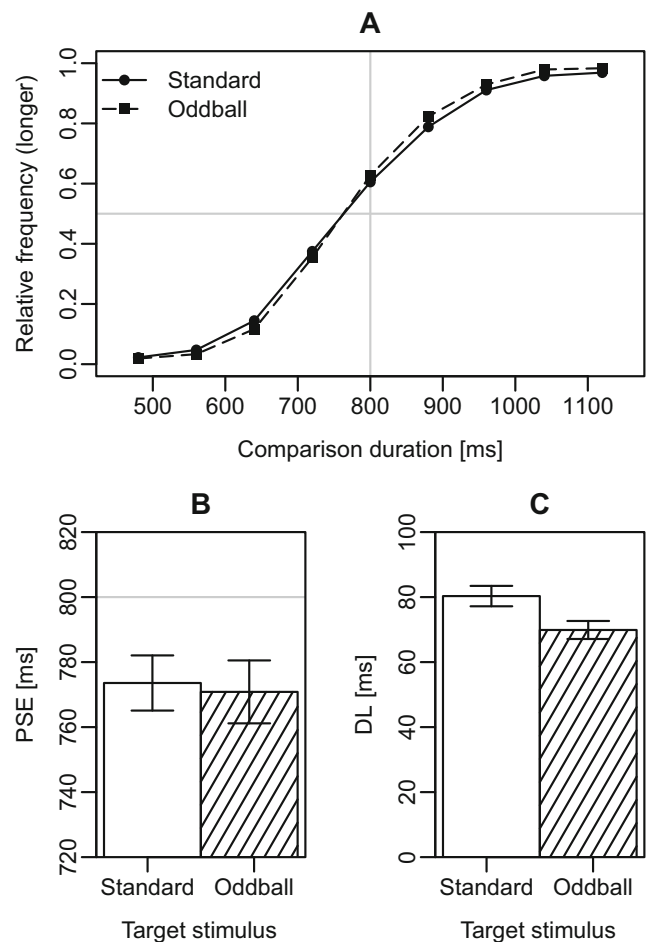

Fig. 4 A Relative frequencies for longer responses as a function of comparison duration and target stimulus in Experiment 2. Circles (squares) indicate relative frequencies in the case of a standard (oddball) target, averaged across participants. The horizontal gray line indicates the probability for a longer judgment of 0.5. The vertical gray line indicates the objective standard duration. **B** Point of subjective equality (PSE) as a function of target stimulus. **C** Difference limen (DL) for standard and oddball targets; DLs were calculated individually and are depicted averaged across participants. All error bars indicate ±1 standard error

Results

Point of subjective equality Figure 4 depicts relative frequencies of longer responses (Panel A) and PSEs for both conditions (Panel B). A one-factorial repeated measures ANOVA on PSE revealed no effect of target stimulus on PSE, $F < 1$.[3] The mean PSE for standard targets (774 ms)

---

[3] A power analysis revealed that with an alpha level of .05, a SOE of 25 ms, and a sampling variability of 38 ms (resembling the standard deviation of the SOEs in this experiment), the statistical power of obtaining a significant result is larger than 99 %.

was not significantly different from the mean PSE for oddball targets (771 ms), thus no indication of a SOE was present. Consistent with Experiment 1, separate one-sample *t*-tests revealed that both PSEs were significantly smaller than the objective standard duration; $t(39) = 3.11$, $p = .003$, for the standards and $t(39) = 3.01$, $p = .005$, for the oddballs.

Difference limen Mean DLs for standard and oddball targets are shown in Panel C of Fig. 4. The repeated measures ANOVA revealed a significant effect of target stimulus, $F(1, 39) = 16.51$, $MSE = 132$, $p < .001$, $\eta_p^2 = .30$, indicating larger DLs for standard targets (80 ms) than for oddball targets (70 ms). Replicating the results of Experiment 1, discrimination sensitivity was again higher when oddballs had to be judged than when standards had to be judged.

Response time Figure 5 displays the mean RTs for standard and oddball targets as a function of comparison duration. The repeated measures ANOVA on RT revealed a significant effect of target stimulus, $F(1, 39) = 19.41$, $MSE = 2,769$, $p < .001$, $\eta_p^2 = .33$, indicating generally faster responses to oddball targets (733 ms) than to standard targets (785 ms).

Point of maximal uncertainty Mean PMUs were 775 ms, irrespective of target stimulus (grand mean). As in Experiment 1, a one-sample *t*-test showed that the mean PMU was significantly smaller than the objective standard duration of 800 ms, $t(39) = 9.01$, $p < .001$, which is again in line with the PSE results.

Catch Trials Catch stimuli were successfully detected in 99.8 % of all catch trials in Experiment 2, thus confirming that participants attended to the stimuli. The mean RT in catch trials was 955 ms.

Discussion

The results of Experiment 2 replicated those of Experiment 1. First, subjective durations for oddball and standard targets did not differ. The lack of a SOE in this experiment was especially surprising as the occurrence probability of oddballs was further reduced. The probability for each of the nine possible oddballs in the target position was only 3.3 %, and the probability for any oddball was only 30 %. Second, as in Experiment 1, participants overestimated the duration of both standard and oddball targets compared to the objective standard duration. Third, a higher discrimination sensitivity for oddballs than for standards was again observed. Finally, RTs to oddball targets were shorter than RTs to standard targets, whereas the PMUs again mirrored the PSE results.

Experiment 3

In Experiments 1 and 2 a clockwise arrangement of stimuli highlighted the temporal position of the target stimulus. While this modification enabled target identification without the need to count the stimuli, the procedure also deviated from the classic oddball paradigm in which all stimuli are presented in the same spatial position. To investigate whether spatial arrangement influences perceived duration, we conducted another experiment. In Experiment 3, all stimuli were presented at the same spatial position in the center of the screen and a clockwise moving orientation point indicated the temporal position of the target stimulus.

Method

Participants A fresh sample of forty members of the University of Tübingen (28 female, 18 - 47 years, $M = 23.0$
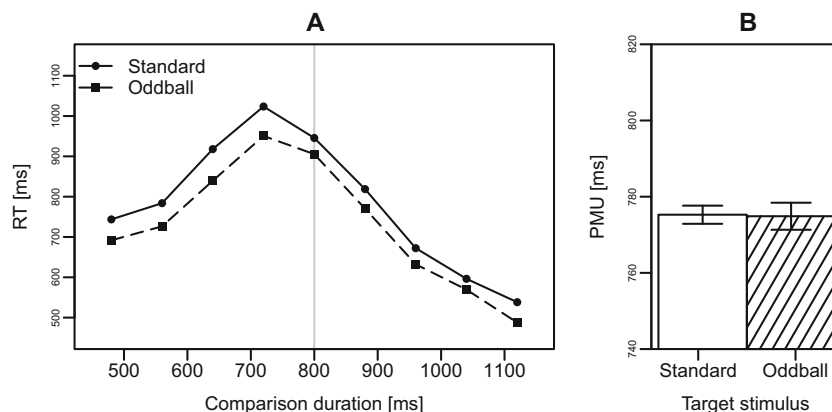

Fig. 5 A Mean response times (RT) for standard and oddball targets in Experiment 2 as a function of comparison duration. The vertical gray line indicates the objective standard duration of 800 ms. **B** Point of maximal uncertainty (PMU) as a function of target stimulus. Error bars indicate ±1 standard error

years) participated in the experiment. Thirty-seven of them were right-handed; all had normal or corrected-to-normal vision and normal color vision. The experiment lasted approximately 90 minutes and participants either received partial course credit or were paid € 12. One additional student participated, but his data had to be excluded due to more than five missed catch trials.

Apparatus, stimuli, and durations A MAC computer controlled stimulus presentation and recorded the participants' responses. The same screen was used as in Experiment 1 and 2. The identical stimuli and durations were used as in Experiment 2.

Procedure After a practice block of ten trials, four experimental blocks of 100 trials each followed. As in Experiment 2, targets were standard stimuli in 70 % of the experimental trials and oddball stimuli in 30 % of the experimental trials. These 30 % oddball targets were distributed across the three oddball shapes (balanced) and three oddball colors (randomized). The difference between Experiments 3 and 2 was that stimuli were no longer arranged in a clockwise manner but were all presented in the center of the screen, while the orientation point moved around the stimuli (see Fig. 1B). Each trial started with the presentation of the orientation point at the 6 o'clock position and after an ISI, the first standard was presented in the middle of the screen with the orientation point below the standard. Then with the onset of every next standard presentation, the orientation point jumped clockwise to the next position (9 o'clock with the second standard, 12 o'clock with the third standard, 3 o'clock with the fourth standard). Finally, the orientation point was presented below the target stimulus (6 o'clock position). The orientation point remained in its location during any ISI, so it was always present in the screen just as in Experiments 1 and 2. Note that the spatial relationship between the stimuli and the orientation point was therefore identical to Experiments 1 and 2. Participants were instructed to fixate the centrally presented stimuli and to use the (peripheral) orientation point only in order to anticipate the target stimulus in the sequence. Catch trials and breaks were the same as in Experiment 2.

Design and data analysis This experiment was a one-factorial design with the within-subject factor Target Stimulus (standard target vs. oddball target). PSEs, SOEs, DLs, and PMUs were calculated analogously to Experiments 1 and 2.

Results

Point of subjective equality Figure 6 depicts relative frequencies of longer responses (Panel A) and PSEs for both
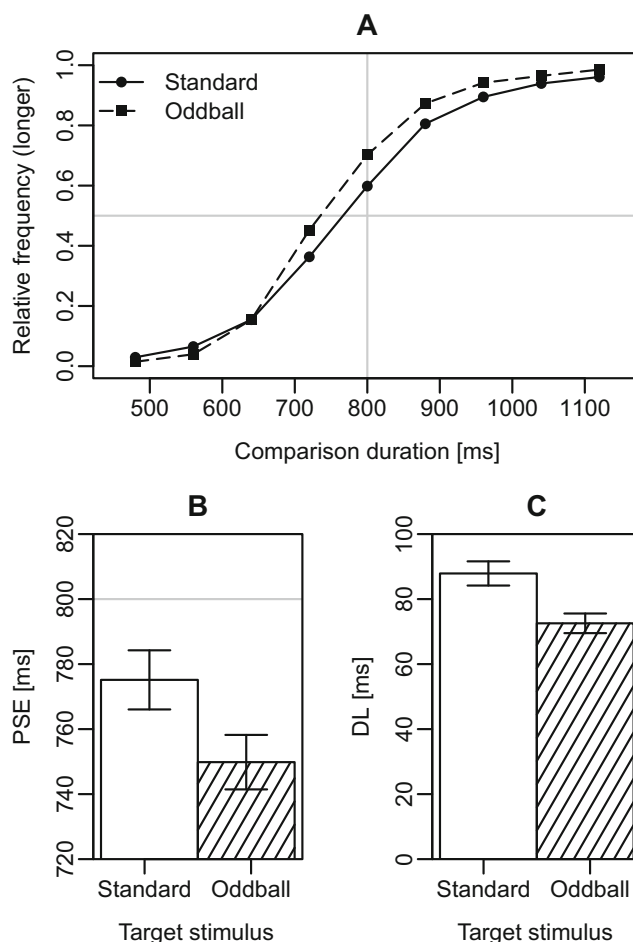

Fig. 6 A Relative frequencies for longer responses as a function of comparison duration and target stimulus in Experiment 3. Circles (squares) indicate relative frequencies in the case of a standard (oddball) target, averaged across participants. The horizontal gray line indicates the probability for a longer judgment of 0.5 and the vertical gray line indicates the objective standard duration. **B** Point of subjective equality (PSE) as a function of target stimulus. **C** Difference limen (DL) for standard and oddball targets; DLs were calculated individually and are depicted averaged across participants. All error bars indicate ±1 standard error

conditions (Panel B). A one-factorial repeated measures ANOVA on PSE revealed a significant effect of target stimulus on PSE, $F(1, 39) = 20.75$, $MSE = 618$, $p < .001$, $\eta_p^2 = .35$. The mean PSE for oddball targets (750 ms) was significantly smaller than for standard targets (775 ms), thus indicating a SOE of about 25 ms. As in Experiments 1 and 2, separate one-sample t-tests showed that both PSEs were significantly smaller than the objective standard duration; $t(39) = 2.73$, $p = .009$ for the standards and $t(39) = 5.99$, $p < .001$ for the oddballs. Thus, all targets were still overestimated in this experiment, but this time oddball targets were even more overestimated than standard targets.

Difference limen Mean DLs for standard and oddball targets are shown in Panel C of Fig. 6. The repeated measures ANOVA showed a significant effect of target stimulus, $F(1, 39) = 36.34$, $MSE = 130$, $p < .001$, $\eta_p^2 = .48$, indicating larger DLs for standard targets (88 ms) than for oddball targets (73 ms). As in both previous experiments, discrimination sensitivity was better for oddballs than for standards.

Response time Figure 7 displays the mean RTs for standard and oddball targets as a function of comparison duration. The corresponding repeated measures ANOVA again showed a significant effect of target stimulus on RT, $F(1, 39) = 27.81$, $MSE = 4,358$, $p < .001$, $\eta_p^2 = .42$, indicating generally faster responses to oddball targets (812 ms) than to standard targets (890 ms).

Point of maximal uncertainty Mean PMUs were 776 ms for standard targets and 774 ms for oddball targets. A repeated measures ANOVA showed that the difference between these two PMU values was not statistically meaningful, $F < 1$. Another one-sample t-test showed that the grand mean of PMU was significantly smaller than the objective standard duration, $t(39) = 10.50$, $p < .001$. In contrast to Experiments 1 and 2, PMU results did not fully mimic the PSE results.

Catch Trials Catch stimuli were successfully detected in 99.6 % of all catch trials in Experiment 3. The mean RT in catch trials was 861 *ms*.

Discussion

Experiment 3 investigated whether the spatial arrangement of stimuli influences perceived duration. Consistent with the previous experiments, the duration of target stimuli was generally overestimated and discrimination sensitivity was superior for oddball targets as compared to standard targets. Furthermore, responses associated with oddball targets were faster than responses associated with standard targets, and PMUs were located to the left of the standard duration, strengthening the view that participants experienced the largest level of uncertainty when target stimuli were slightly shorter than standards. However, in contrast to the previous experiments, the temporal overestimation was even more pronounced for oddballs than for standards, that is, a genuine oddball effect occurred in Experiment 3.

General discussion

This study investigated the temporal oddball effect within two modified versions of the classic stream-based oddball paradigm (Birngruber et al. 2014a; Chen and Yeh 2009; Kim and McAuley 2013; New and Scholl 2009; Pariyadath and Eagleman 2007; Schindel et al. 2011; Tse et al. 2004). In contrast to the classic paradigm, the present versions required duration judgments not only about oddballs but also about standards, thereby enabling a comparison of these two judgments. If oddballs were generally perceived as being longer than standards, one would expect a meaningful difference between these two measures.

In three experiments, the occurrence probability of oddballs and the spatial arrangement of the stimuli was manipulated. In all experiments, the duration of targets was generally overestimated in comparison to the standards presented in the stream. In Experiments 1 and 2, this overestimation was of a similar size for standard and oddball targets. Only in Experiment 3, in which all stimuli were presented at the same spatial position, was an additional overestimation of oddball targets as compared to standard targets and thus a genuine oddball effect observed. The general
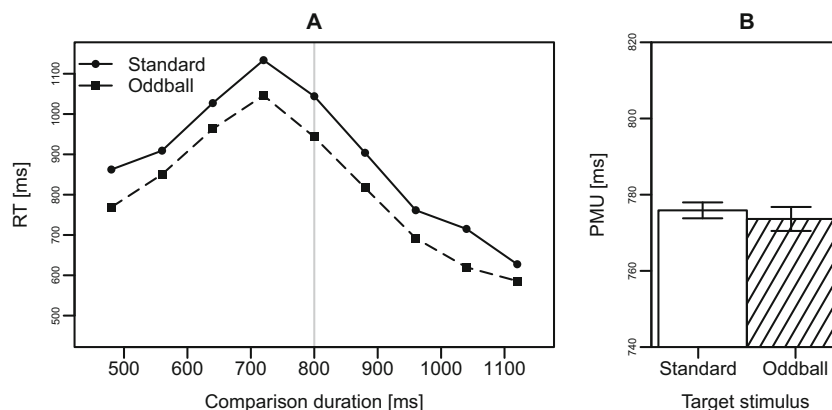

Fig. 7 A Mean response times (RT) for standard and oddball targets in Experiment 3 as a function of comparison duration. The vertical gray line indicates the objective standard duration of 800 ms. **B** Point of maximal uncertainty (PMU) as a function of target stimulus. Error bars indicate ±1 standard error

overestimation of all target stimuli was also reflected in the newly introduced measurement of judgment uncertainty, the *Point of maximal uncertainty* (PMU). This measurement illustrated that the longest RTs occurred for targets that were shorter than the standard duration, irrespective of whether the target was a standard or an oddball, indicating that participants experienced maximal uncertainty when target durations were slightly shorter than standard durations.

The general overestimation could be the result of a negative time-order error because in the present experiments (and most previous oddball studies), target stimuli were presented in late positions of the stream and thus had to be compared to earlier presentations of the standard. The negative time-order error is usually explained with the fading-trace theory. This theory assumes that the memory of the first presented stimulus starts to fade away before the second stimulus is presented (Schab and Crowder 1988; Köhler 1923). Thus, it is proposed that the fading representation of the first stimulus results in a shortened subjective duration for this stimulus and therefore an overestimation of a later presented stimulus.

Importantly, such a negative time-order error could have at least inflated the oddball effects reported in previous studies, because usually only the duration of oddballs was judged by the participants. The finding of a general overestimation of all targets demonstrates the importance of assessing perceived duration of both oddballs and standards.

In the present experiments, the frequency of oddballs was manipulated because it is generally assumed that stimulus probability effectively alters perceived duration (Kim and McAuley 2013; Schindel et al. 2011; Ulrich et al. 2006). According to this idea, stimulus probability affects stimulus expectancy and unexpected stimuli are coded less efficiently. This may result in a larger neural response for infrequent and thus unexpected stimuli as compared to frequent, expected ones (Gotts et al. 2012; Henson 2003; Summerfield et al. 2008; Wiggs and Martin 1998). Because duration estimation is assumed to reflect the size of the neural response (Eagleman and Pariyadath 2009; Pariyadath and Eagleman 2007, 2012), infrequent stimuli would be temporally overestimated. In line with this notion, we observed longer judged durations for infrequent oddball targets than for frequent standard targets in Experiment 3. This result illustrates that subjective overestimations of oddballs do exist beyond the general overestimation effect of all targets (see also Ulrich et al. 2006 and van Wassenhove et al. 2008).

However, no such subjective oddball effect occurred in Experiments 1 and 2. The absence of a subjective oddball effect in Experiment 2 is especially surprising, because in this experiment the same frequency manipulation was employed as in Experiment 3. Experiments 1 and 2 differed from Experiment 3 only in regard to the spatial presentation of stimuli. Specifically, stimuli were presented in a clockwise arrangement in Experiments 1 and 2 whereas all stimuli were presented centrally in Experiment 3 (and most previous studies, e.g., Tse et al. 2004 and van Wassenhove et al. 2008). It is unclear how differences in spatial arrangement could account for the differential occurrence of a temporal oddball effect, because this effect is commonly understood as mainly feature-based rather than spatially-specific (see New and Scholl 2009). Next we will consider possible explanations for the unexpected absence of a subjective oddball effect in Experiments 1 and 2.

First, in these experiments participants had to make saccades in order to follow the stream of stimuli and it has been argued that saccades can influence duration perception (Morrone et al. 2005; Yarrow et al. 2001). However, such an influence should have affected the perception of standard and oddball targets equally and thus cannot easily account for the lack of a subjective oddball effect. Second, according to the repetition suppression hypothesis (Gotts et al. 2012; Grill-Spector et al. 2006), the oddball effect arises because of a reduced neural response to repeated standards. This reduction might be less pronounced for repeated stimuli being presented in different spatial locations resulting in a similar neural response for both standard and oddball targets in Experiments 1 and 2. Finally, it can be speculated that repeated stimuli in the same position are treated as a single object while repeated stimuli in different locations are represented as distinct objects. This might reduce the unexpectedness of oddball targets in Experiments 1 and 2. Clearly, future research is necessary to uncover the underlying mechanisms by which spatial features influence the occurrence of temporal oddball effects.

In summary, the present study demonstrates that the duration of both standard and oddball stimuli are overestimated in stream-based oddball paradigms. This finding highlights the importance of comparing the judged duration of oddballs and standards in order to evaluate the net effect of oddness. Furthermore, the present study suggests that the occurrence of temporal oddball effects depends on spatial features of stimulus presentation.

References

Allan, L.G. (1977). The time-order error in judgments of duration. *Canadian Journal of Psychology*, *31*, 24–31.

Birngruber, T., Schröter, H., & Ulrich, R. (2014a). Duration perception of visual and auditory oddball stimuli: Does judgment task modulate the temporal oddball effect? *Attention, Perception, & Psychophysics*, *76*, 814–828.

Birngruber, T., Schröter, H., & Ulrich, R. (2014b). What makes an oddball odd? Evidence from a spatially predictable temporal oddball paradigm. *Procedia-Social and Behavioral Sciences*, *126*, 190–191.

Birren, J.E., & Botwinick, J. (1955). Speed of response as a function of perceptual difficulty and age. *Journal of Gerontology*, *10*, 433–436.

Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Cacioppo, J.T., & Dorfman, D.D. (1987). Wave-form moment analysis in psychophysiological research. *Psychological Bulletin*, *102*, 421–438.

Chen, K.M., & Yeh, S.L. (2009). Asymmetric cross-modal effects in time perception. *Acta Psychologica*, *130*, 225–234.

Dyjas, O., & Ulrich, R. (2014). Effects of stimulus order on discrimination processes in comparative and equality judgements: Data and models. *The Quarterly Journal of Experimental Psychology*, *67*, 1121–1150.

Eagleman, D.M., & Pariyadath, V. (2009). Is subjective duration a signature of coding efficiency? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 1841–1851.

Gotts, S.J., Chow, C.C., & Martin, A. (2012). Repetition priming and repetition suppression: A case for enhanced efficiency through neural synchronization. *Cognitive Neuroscience*, *3*, 227–237.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*, 14–23.

Grondin, S., Ouellet, B., & Roussel, M.E. (2004). Benefits and limits of explicit counting for discriminating temporal intervals. *Canadian Journal of Experimental Psychology*, *58*, 1–12.

Guilford, J.P. (1954). *Psychometric methods*. New York: McGraw-Hill.

Henmon, V.A.C. (1910). The relation of the time of a judgment to its accuracy. *Psychological Review*, *18*, 186–201.

Henson, R.N.A. (2003). Neuroimaging studies of priming. *Progress in Neurobiology*, *70*, 53–81.

Jamieson, D.G., & Petrusic, W.M. (1975). Presentation order effects in duration discrimination. *Perception & Psychophysics*, *17*, 197–202.

Kim, E., & McAuley, J.D. (2013). Effects of pitch distance and likelihood on the perceived duration of deviant auditory events. *Attention, Perception, & Psychophysics*, *75*, 1547–1558.

Köhler, W. (1923). Zur Theorie des Sukzessivvergleichs und der Zeitfehler. *Psychological Research*, *4*, 115–175.

Lapid, E., Ulrich, R., & Rammsayer, T. (2008). On estimating the difference limen in duration discrimination tasks: A comparison of the 2AFC and the reminder task. *Perception & Psychophysics*, *70*, 291–305.

Matthews, W.J. (2011). Stimulus repetition and the perception of time: The effects of prior exposure on temporal discrimination, judgment, and production. *PLoS ONE*, *6*, e19815.

Miller, J., & Ulrich, R. (2001). On the analysis of psychometric functions: The Spearman-Kärber method. *Perception & Psychophysics*, *63*, 1399–1420.

Morrone, M.C., Ross, J., & Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. *Nature Neuroscience*, *8*, 950–954.

New, J.J., & Scholl, B.J. (2009). Subjective time dilation: Spatially local, object-based, or a global visual experience *Journal of Vision*, *9*, 1–11.

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

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

Pelli, D.G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Rattat, A.C., & Droit-Volet, S. (2012). What is the best and easiest method of preventing counting in different temporal tasks? *Behavior Research Methods*, *44*, 67–80.

Schab, F.R., & Crowder, R.G. (1988). The role of succession in temporal cognition: Is the time-order error a recency effect of memory? *Perception & Psychophysics*, *44*, 233–242.

Schindel, R., Rowlands, J., & Arnold, D.H. (2011). The oddball effect: Perceived duration and predictive coding. *Journal of Vision*, *11*, 1–9.

Seifried, T., & Ulrich, R. (2010). Does the asymmetry effect inflate the temporal expansion of odd stimuli? *Psychological Research*, *74*, 90–98.

Smith, P.L., & Vickers, D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology*, *32*, 135–168.

Summerfield, C., Trittschuh, E.H., Monti, J.M., Mesulam, M.M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, *11*, 1004–1006.

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

Ulrich, R., Nitschke, J., & Rammsayer, T. (2006). Perceived duration of expected and unexpected stimuli. *Psychological Research*, *70*, 77–87.

Ulrich, R., Wing, A.M., & Rinkenauer, G. (1995). Amplitude and duration scaling of brief isometric force pulses. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1457–1472.

Vickers, D., & Packer, J. (1982). Effects of alternating set for speed of accuracy on response time, accuracy and confidence in a unidimensional discrimination task. *Acta Psychologica*, *50*, 179–197.

Van Wassenhove, V., Buonomano, D.V., Shimojo, S., & Shams, L. (2008). Distortions of subjective time perception within and across senses. *PLoS ONE*, *3*, e1437.

Wiggs, C.L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology*, *8*, 227–33.

Woodworth, R., & Schlossberg, H. (1965). *Experimental Psychology (3. revised éd.)* New York: Holt, Rinehart and Winstin.

Yarrow, K., Haggard, P., Heal, R., Brown, P., & Rothwell, J.C. (2001). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. *Nature*, *414*, 302–305.