Sustained Attention Ability Affects Simple Picture Naming

Suzanne R. Jongman

Sustained attention has previously been shown as a requirement for language production. However, this is mostly evident for difficult conditions, such as a dual-task situation. The current study provides corroborating evidence that this relationship holds even for simple picture naming. Sustained attention ability, indexed both by participants’ reaction times and individuals’ hit rate (the proportion of correctly detected targets) on a digit discrimination task, correlated with picture naming latencies. Individuals with poor sustained attention were consistently slower and their RT distributions were more positively skewed when naming pictures compared to individuals with better sustained attention. Additionally, the need to sustain attention was manipulated by changing the speed of stimulus presentation. Research has suggested that fast event rates tax sustained attention resources to a larger degree than slow event rates. However, in this study the fast event rate did not result in increased difficulty, neither for the picture naming task nor for the sustained attention task. Instead, the results point to a speed-accuracy trade-off in the sustained attention task (lower accuracy but faster responses in the fast than in the slow event rate), and to a benefit for faster rates in the picture naming task (shorter naming latencies with no difference in accuracy). Performance on both tasks was largely comparable, supporting previous findings that sustained attention is called upon during language production.

Keywords: language production; sustained attention; object naming; event rate; individual differences

Introduction

Accumulating evidence suggests that word production requires attention even though it is a highly practiced skill. For instance, language production has been shown to impair performance on an unrelated task such as driving a car when speaking and driving co-occur (Kubose et al., 2006). Several other studies have shown similar effects and argued that word production draws from a central attentional system (Cook & Meyer, 2008; Ferreira & Pashler, 2002; Roelofs, 2008). However, it remains unclear which attention system these studies refer to. Attention is a broad term that comprises different functionally and anatomically separate subsystems. Posner and colleagues have proposed there are three such attention systems: executive control, orienting, and alerting (Petersen & Posner, 2012; Posner & Petersen, 1990; Posner & Rothbart, 2007). One or all of these systems could contribute to language production, possibly in different ways.

Executive control is needed to keep one’s goals in mind to successfully complete an action. This first attention system has been studied in relation to language, mostly in comprehension (Ye & Zhou, 2009) but also recently in production (Shao, Roelofs, & Meyer, 2012). Shao et al. decomposed executive control into three subcomponents – updating, inhibiting, and shifting – as suggested by Miyake and colleagues (Miyake et al., 2000). Updating (maintaining and manipulating items in working memory), and inhibiting (suppressing an inappropriate response) were found to correlate with participants’ speed of word production. However, the third component, shifting (switching between goals) did not show such a correlation. This suggests that some but not all types of executive attention play a role during word production.

Second, orienting is the system that moves attention towards new and relevant information. Finally, the alerting system is needed to heighten levels of attention. This can be a short-lived attention boost for instance after a warning signal, or a prolonged attention increase for instance during a task. Such prolonged maintenance of attention is now referred to as sustained attention and was previously known as vigilance. Sustained attention has been suggested to play a role during word production, based on results of studies using the individual differences approach (Jongman, Meyer, & Roelofs, 2015; Jongman, Roelofs, & Meyer, 2015). In three experiments we showed that poor sustained attention ability coincided with worse performance on a picture naming task such that naming latency distributions were more positively skewed (i.e. had more ‘abnormally’ slow trials). These studies show there is
a consistent relationship between sustained attention and language production.

One problem with these previous studies is that the strongest correlations were found for picture naming in a relatively difficult setting, i.e. in dual-task situations where participants named pictures and concurrently performed a second task, either linguistic or non-linguistic. Significantly smaller correlations were found for simple picture naming. One aim of the present study is to see whether the relationship between sustained attention and simple picture naming is reliably present. More importantly, the problem with these previous studies is that they provide only correlational evidence. These individual differences studies point to a role of sustained attention during language production, but as of yet there is no definitive evidence that sustained attention is required to produce words. To show a causal link I manipulated the need for sustained attention in a picture naming task by using a variable that is known to tax sustained attention in traditional sustained attention tasks. This manipulation could provide more direct evidence that sustained attention is required for fluent language production.

Sustained attention was first studied by Mackworth during the second world war. During the war, radar and sonar operators needed to detect rare irregular events for many hours. Mackworth showed that as time progressed, the likelihood of missing such a rare target would increase (Mackworth, 1948). Since then it has been shown that sustained attention can be affected by three factors: task parameters, participant characteristics, and environmental conditions (Ballard, 2001; Langner & Eickhoff, 2013; Oken, Salinsky, & Elsas, 2006; Robertson & O’Connell, 2010; Sarter, Givens, & Bruno, 2001). Task parameters that tax sustained attention include infrequent target signals, degraded stimuli, spatial uncertainty and high speed of stimulus presentation (McFarland & Halcomb, 1970; Mouloua & Parasuraman, 1995; Parasuraman, 1979; Parasuraman, Nestor, & Greenwood, 1989). The second factor that can influence sustained attention performance relates to individuals’ characteristics. Older participants have more difficulty with maintaining attention than younger adults (McFarland & Halcomb, 1970; Mouloua & Parasuraman, 1995; Parasuraman, 1979; Parasuraman, Nestor, & Greenwood, 1989). Some clinical populations appear to have deficits in sustained attention, such as persons with schizophrenia and individuals with ADHD (Epstein, Johnson, Varia, & Conners, 2001; Liu et al., 2002). Finally, environmental factors such as noise can affect sustained attention performance (Broadbent & Gregory, 1965).

In the present study one of the variables known to tax sustained attention was manipulated, namely the speed of stimulus presentation. The speed of stimulus presentation can be manipulated either by changing the inter-stimulus interval (ISI) or by changing the duration of the stimulus itself. Jerison and Pickett (1964) first reported a threefold decrease in the hit rate when the event rate in a visual vigilance task was increased from 5 events per minute to 30 events per minute. In other words, there are more failures to detect a target when the stimuli are presented in rapid succession. Other studies have also observed lower hit rates for faster event rates (1964). Some studies have, in addition, found a greater vigilance decrement (worse performance as time on task increases) for higher event rates compared to slow event rates (1964). The fact that fast event rates cause worse performance on sustained attention tasks has been argued to be due to faster depletion of a limited pool of attentional resources (Warm, Parasuraman, & Matthews, 2008).

These mentioned studies have all mainly looked at accuracy, not reaction time (RT). Ballard (2001) did measure RTs and found that participants responded faster in a fast event rate as compared to slow event rates. The decrease both in hit rate and in RTs for fast compared to slow event rates could indicate a speed-accuracy trade-off instead of reflecting sustained attention depletion differences between fast and slow tasks. However, findings that false alarms, failures to withhold response to a non-target, are higher for slow event rates than for fast rates argue against this idea of a speed-accuracy trade-off (Koelega et al., 1992; Lanzetta et al., 1987). Moreover, a large vigilance decrement in the fast event rate in both hit rate (a decrease over time) and RTs (an increase over time) would also argue against such a trade-off.

There have been several studies reporting contrasting results, such that fast event rates actually cause better performance than slow rates. This is true for sustained attention studies with children (Chee, Logan, Schachar, Lindsay, & Wachsmuth, 1989; Rose, Murphy, Schickedanz, & Tucci, 2001) but has also been found in a study using adults, albeit in an experiment not designed to test sustained attention. De Jong, Berendsen, and Cools (1999) had participants perform a spatial version of the Stroop task (i.e. the word LOW/HIGH presented above/below the center, with participants responding to location only), either in a fast or in a slow event rate design. Participants were faster to respond for the fast event rate with no loss in accuracy levels, as compared to the slow event rate. Interestingly, there was a large reduction in the Stroop effect (slower responses to incongruent trials, such as LOW presented above the center, than to congruent trials) in the fast event rate. The authors argued that with a fast event rate attention was sharply focused on the task, resulting in fast responses and fewer opportunities for the word meaning to interfere.

One goal of the present study was to provide a better picture of whether differences in the degree of depletion of sustained attention account for effects of event rate, whether it is a mere speed-accuracy trade-off or whether event rate differences reflect variation in the focusing of attention. Participants performed a digit discrimination task (DDT) measuring sustained attention with a fast and slow event rate. The DDT is a visual continuous performance task in which digits are presented on the screen one by one and participants are instructed to respond to an infrequent target only, i.e. the digit zero among foils one to nine (Matthews & Davies, 2001; Parasuraman et al., 1989; Sepede et al., 2012). In the present experiment, the ISI between digits was manipulated to create two contrasting event rates:
fast versus slow. Stimulus duration was held constant for both conditions. The analyses not only focused on hit rate and false alarms but also on RTs. People should be faster to respond in the fast event rate, as shown in previous studies. If there is no loss in accuracy then the findings are in accordance with the sharp focusing of attention in the fast rate as in De Jong and colleagues (Matthews & Davies, 2001; Parasuraman et al., 1989; Sepehri et al., 2012). If however, hit rates are lower for the fast event rate condition as compared to the slow event rate this would not hold. If fast event rates indeed tax sustained attention to a larger degree, a larger decrement should be found in both hit rates and RTs such that participants have fewer hits and respond more slowly over time as compared to the slow event rate. Lack of such an increased decrement would point to a mere speed-accuracy trade-off.

The main goal of this study however, was to show that word production requires sustained attention. Participants not only performed a sustained attention task, but also a picture naming task. The same manipulation known to tax sustained attention in traditional sustained attention tasks was used in the picture naming task: pictures were presented either in a fast or slow event rate. In both event rates, pictures of simple objects were presented for one second and participants were asked to name them. In the fast event rate pictures were separated by a blank screen for only 500 ms, whereas the ISI was 2000 ms in the slow event rate condition. If sustained attention is required for word production one would expect patterns in error rates and RTs for the two event rates comparable to the pure sustained attention task. So if the fast event rate in the picture naming task depletes attention to a larger degree, one should find more errors and a larger decrement over time in both errors and RTs as compared to the slow naming task. Conversely, if the fast event rate helps to maintain focus on the task at hand, RTs should be lower in the fast rate and accuracy levels should be the same or higher when compared to the slow event rate.

Moreover, if sustained attention is required for picture naming, word production performance should correlate with performance on the pure sustained attention task. Thus individuals who are better at the sustained attention task should also be better at the picture naming task. In our previous studies accuracy in the sustained attention tasks and language production tasks was very high, and sustained attention ability was always defined by RTs. Manipulating event rate should result in more errors and allow sustained attention ability to be defined not only by RTs but also by hit accuracy.

Mean RTs on the picture naming task were divided into two dissociable components to test whether event rate affected all of the responses or only a subset. RT distributions are often not normally distributed but positively skewed. Instead of transforming the distribution, one can use ex-Gaussian analysis to decompose the distribution to get a better understanding of individuals’ behavior. The normal part of the distribution is indexed by the $\mu$ parameter, whereas the tail end (the abnormally slow responses) is indexed by $\tau$. In previous experiments we found sustained attention ability to correlate with the $\tau$ parameter of picture description latencies, but not with $\mu$ (Jongman, Meyer, & Roelofs, 2015; Jongman, Roelofs, & Meyer, 2015). In other words, individuals with poorer sustained attention ability did not have an overall right shift in naming latency distributions, but did have a larger right tail (i.e. more instances of a very slow response). We took $\tau$ to reflect the instances where individuals experience a lapse of attention, which has been previously proposed by Unsworth et al. (2010). It could very well be the case that the effect of event rate is manifested mostly in the $\tau$ parameter and as such correlations could be stronger for $\tau$ as compared to $\mu$.

In summary, the first aim of the present study was to perform the first causal test of sustained attention in word production by manipulating the need for sustained attention within a language production task by varying event rate, and compare it to the effect of the same manipulation in a traditional sustained attention task. Does the event rate manipulation cause similar effects in both tasks? The second aim of this study was to find out if differences between a fast and slow event rate are due to differences in attention depletion, adjustments in focusing of attention, or reflect a speed-accuracy trade-off? The third aim of this study was to extend previous individual differences studies on sustained attention and language production. Does sustained attention ability, as measured not only by RTs but also by accuracy, correlate with simple picture naming?

### Method

#### Participants

Forty-eight young adults participated in the experiment; all were students of the Radboud University Nijmegen or the Hogeschool van Arnhem en Nijmegen. All subjects were native speakers of Dutch and had normal or corrected-to-normal vision. The average age was 23.0 years (range: 19–32), forty-one participants were female, forty-three were right-handed. Ethical approval was granted by the Ethics Board of the Faculty of Social Sciences of the Radboud University. Participants provided written informed consent before starting the experiment, participants received monetary compensation.

#### General Procedure

Participants were tested individually in a dimly illuminated, soundproof booth. The tasks were presented on a 17 inch (iiyama LM704UT) screen, using Presentation Software (Version 16.2, www.neurobs.com). Participants first performed the picture naming task, with two fast event rate blocks alternating with two slow event rate blocks. Participants then performed the digit discrimination task, again with four blocks in total, alternating between fast and slow event rates.

#### Picture Naming Task

**Materials and design**

Participants were presented with thirty common objects, with each object shown thirty times. The object names were monosyllabic and highly frequent (mean lemma frequency: 600 tokens per million; CELEX database...
The pictures were selected for high name agreement with a mean of 89% (Severens et al., 2005). See supplementary file S1 for all object names.

The pictures were depicted on the center of the computer screen, 300 by 300 pixels, corresponding to visual angles of 7.0° horizontally and 6.4° vertically when viewed from the participant’s position, approximately 60 cm away from the screen. The thirty pictures were presented in a set, in a pseudorandomized order such that two objects of the same semantic category never followed one another, nor did two names starting with the same phoneme.

Procedure
Prior to the start of the experiment, participants were shown the pictures together with the corresponding names. In the first familiarization block, each picture was presented in the middle of the screen with its name written below. The participant pressed Enter to proceed to the next picture. In the second familiarization block, a picture was presented and participants were asked to name the picture. Once the voicekey was triggered the correct name was shown on the screen, and participants were asked to check if their response matched the written text. Once all thirty pictures had been named, the participant proceeded with the actual experiment.

In the fast event rate, a trial was initiated by a blank screen shown for 500 ms, then the picture was presented for 1000 ms. In the slow event rate, the blank screen initializing the trial was presented for 2000 ms. The duration of picture presentation was identical to the fast event rate, thus 1000 ms. These durations were chosen based on the taxonomy suggested by Parasuraman and Davies (1977), and Lanzetta and colleagues (1987). The first study suggested a cut-off of 24 event rates per minute as the transition from a slow event rate to a fast event rate, whereas the latter study suggested a higher cut-off such as 48 event rates per minute for a simple task (i.e. when the current trial does not depend on information from the previous trial). Here, 40 events were presented per minute in the fast event rate, chosen to be near the Lanzetta cut-off whilst still allowing for enough time to name each picture before the start of the next picture onset.

In the fast event rate a total of 600 pictures were presented, divided over two blocks. The slow event rate consisted of 300 pictures in total, also over two blocks. Fast and slow event rate had equal durations, namely 15 minutes. All participants were presented with both event rates, thus with four blocks. Fast and slow event rate blocks alternated, with block order counterbalanced across the participants.

Analyses
Vocal responses were recorded by a microphone (Sennheiser ME64). Response onsets were automatically marked in Praat (Boersma & Weenink, 2012) using a personalized script for each participant. Hesitations and naming errors were removed from the analyses. Furthermore, a reliability check was performed by manually annotating 200 trials randomly selected from 4 randomly chosen participants and comparing them with the RTs as measured by Praat. The automatic onset marker measured RTs in a highly comparable manner to manual annotation, the intraclass-coefficient (ICC) was .86 and so was the correlation between the two measurements. Therefore, the naming latencies as measured by Praat seem reliable and are used for the following analysis.

The naming latencies were analyzed using R (R Core Team, 2012), specifically with the lme4 (Bates, Maechler, & Bolker, 2013) and languageR (Baayen, 2011) packages. The initial linear mixed effects model included event rate (fast vs. slow) and block (first block vs. second block) as fixed effects as well as their interaction. Fixed effects were mean-centered. No outlier exclusion procedure was performed, but RTs were log transformed because of positive skewing. Variables were dropped that did not reliably contribute to model fit, models were compared using a likelihood ratio test. The models included both participant and item as random effects: for both factors and their interaction the intercepts and random slopes were included (Barr et al., 2013).

Digit Discrimination Task
Materials and design
The DDT used here was adapted from the task used in Jongman, Meyer and Roelofs (2015). Single digits (font Arial, size 40) were presented in white on a black background. The target digit was the digit 0, and all other digits (1 through 9) were non-targets. The digit 0 was presented on 1/4 of all trials. Stimuli were presented in a pseudorandom sequence such that identical digits never directly followed one another and the target digit was preceded by each non-target an equal number of times. The experiment consisted of 72 practice trials and 1728 experimental trials.

Procedure
Digits were presented for 100 ms each, with an inter-stimulus-interval (ISI) of 500 ms in the fast event rate and an ISI of 2000 ms in the slow event rate. The fast event rate, with 100 events per minute, was far above the suggested cut-off of 48 events per minute as suggested by Lanzetta and colleagues (1987). Participants were instructed to respond to the digit 0 with a button press using their dominant hand. The fast event rate consisted of 1344 trials, divided over two blocks. The slow event rate included 384 trials in total. Both event rate conditions lasted for 13.5 minutes. Participants were thus presented with four blocks where event rate blocks alternated, with block order counterbalanced across the participants.

Analysis
Both RTs and errors were analyzed, with errors divided into misses (failures to respond to targets) and false alarms (responses to non-targets). A logit mixed model was conducted for the hit rates, correct responses to targets (Jaeger, 2008). The model included event rate (fast vs. slow) and block (first vs. second) as fixed effects and their interaction. Fixed effects were mean-centered. The random factor participant was included in the model,
with intercepts and slopes for the main effects and its interaction.

The linear mixed effects model for the correct RTs was computed identically to the model for picture naming (see above). In the fast event rate some responses were made after 600 ms, so after the next digit was already presented. These were coded as correct RTs. A target was never followed by another target, but always by a non-target. The participant therefore never had to respond twice in a row, so even though the participant may still have been in the process of responding when a new digit was presented there was no chance of missing a new target.

Analyses of Individual Differences
For correct trials in the picture naming task, the ex-Gaussian parameters $\mu$, $\sigma$, and $\tau$ were estimated. These first two parameters index the mean ($\mu$) and standard deviation ($\sigma$) of the normal part of the distribution and $\tau$ reflects both the mean and standard deviation of the exponential portion (i.e. the right tail of the distributions in Figure 1). Contrary to the linear mixed effects analyses, latencies were not log-transformed for the ex-Gaussian analyses. The parameters were estimated separately for the fast and slow event rate conditions for each participant, using the continuous maximum-likelihood method by Van Zandt (2000) as implemented in the program QMPE (Heathcote, Brown, & Cousineau, 2004). Only the $\mu$ and $\tau$ parameters were included in the following analyses to keep the number of comparisons low. These parameters were then correlated with two indices of participants’ performance on the DDT, namely mean RTs and hit rates. Correlations were performed for each event rate separately, for example hit rates for the fast event rate condition on the DDT were correlated with the $\mu$ parameter for the slow event rate condition on the picture naming task. Moreover, vigilance decrement was calculated as RTs on the second block minus the first block for both tasks, and then correlated between the two tasks for each event rate. Spearman’s rho is reported as several variables (hit rates and the $\tau$ parameters) were not normally distributed.

Since a total of 18 correlations were tested, the Benjamini-Hochberg correction was used to control for multiple comparisons. Instead of the familywise error rate, the false discovery rate is controlled resulting in greater power than Bonferroni-type procedures (Bender & Lange, 2001; Benjamini & Hochberg, 1995; Benjamini & Yekutieli, 2001; Williams, Jones, & Tukey, 1999). The $p$-values are sorted and ranked in such a way that the smallest value is given rank 1, the second rank 2 and the largest rank $N$. Then, each $p$-value is multiplied by $N$ and divided by its assigned rank. In the present study, this resulted in the first thirteen correlations to be significant after the Benjamini-Hochberg correction, down to an uncorrected $p$-value of .03.

Results
Picture Naming Task
In 1.2% of all trials participants made a naming error (fast event rate: 1.6%; slow: 0.5%), and on 1.0% of the trials they hesitated (fast: 1.2%; slow: 0.4%). Too few errors were made for any further analysis.

The best-fitting linear mixed effects model for correct naming latencies included main effects of event rate ($\beta = 0.03, SE = 0.01, t = 3.97$) and block ($\beta = 0.05, SE = 0.01$, Figure 1: Density plots for the naming latencies of all participants on the picture naming tasks, separate for fast and slow event rates.)
Jongman: Sustained Attention Ability Affects Simple Picture Naming

Art. 17, page 6 of 13

t = 7.78). Removing either event rate or block significantly decreased model fit ($\chi^2 (1) = 13.88, p < .001$ and $\chi^2 (1) = 39.69, p < .001$, respectively). Including the interaction did not improve model fit ($\chi^2 (1) = 0.02, p = .88$). The model revealed that naming latencies differed for the two event rates such that participants were faster to name the pictures in the fast event rate compared to the slow event rate (fast: 692 ms, $SD = 210$; slow: 710 ms, $SD = 201$). The main effect of block revealed that participants were slower in the second block of the experiment as compared to the first block (first: 676 ms, $SD = 183$; second: 720, $SD = 227$), independent of the manipulation of event rate (see Figure 2). A similar analysis, with trial count kept constant in both event rate conditions instead of task duration, revealed the same pattern (see supplementary file S2 for details).

Digit Discrimination Task

False alarms, responding to non-targets, occurred on 0.5% of the non-target trials (fast event rate: 0.5%; slow event rate: 0.7%), precluding any further analysis. Misses, failures to respond to a target, occurred on 4.8% of the target trials. The logit mixed model on hits, the correct responses to targets, revealed a significant effect of event rate ($\beta = 1.85, SE = 0.41, z = 4.47, p < .001$) and block ($\beta = -0.61, SE = 0.22, z = -2.81, p < .01$). The interaction between event rate and block was not significant ($\beta = -0.77, SE = 0.69, z = -1.12, p = .26$). Hit rates were significantly lower for the fast event rate as compared to the slow event rate (94.4% vs 97.9%). Moreover, a vigilance decrement was evident as hit rates decreased over time (first block: 95.9%; second block: 94.4%), see Figure 3.

The linear mixed effects model for correct RTs revealed a significant main effect of event rate ($\beta = 0.09, SE = 0.01, t = 7.26$) and of block ($\beta = 0.04, SE = 0.01, t = 7.39$). Dropping either of these two main effects resulted in worse model fit (event rate: $\chi^2 (1) = 36.12, p < .001$; block: $\chi^2 (1) = 37.09, p < .001$). Including the interaction between event rate and block did not improve model fit ($\chi^2 (1) = 0.03, p = .86$). Participants were faster to respond in the fast event rate condition as compared to the slow event rate condition as compared to the slow event rate (fast: 420 ms, $SD = 82$; slow: 465 ms, $SD = 134$). Performance deteriorated over time, as participants had an average response time of 421 ms ($SD = 92$) in the first half whereas in the second half they responded around 439 ms ($SD = 103$). The lack of an interaction between event rate and block indicates the vigilance decrement was of similar magnitude in both event rates (see Figure 4). See supplementary file S2 for similar analyses for both hit rate and RT, where the number of trials in each condition was kept constant instead of total duration of each condition, again no interactions between event rate and block were found. It must be noted that the main effect of event rate disappeared for hit rate, as hit rates were high for both event rates. It seems time is a critical factor in reducing accuracy levels in the fast event rate.

Individual Differences

Both mean RT and hit rate on the sustained attention task correlated with the $\tau$ parameter of picture naming, for both event rates, see Table 1 and Figure 5. Overall, participants...
who showed worse sustained attention ability, as indexed by both RTs and hit rates, had more abnormally slow picture naming responses than individuals with better sustained attention. Furthermore, the mean RT on the DDT also correlated with the \( \mu \) parameter of the picture naming latencies: those participants with slower reaction times on the sustained attention task were also consistently slower to name pictures. Note that the significance of each of the correlations does not change after removing the three participants with hit rates below .95.

Vigilance decrement on the DDT (mean RT second block – mean RT first block) for the fast event rate did not
correlate with the vigilance decrement for the fast event rate of the picture naming task ($r = \cdot12, p = \cdot41$), neither did the vigilance decrement for the slow event rates on the two tasks ($r = \cdot12, p = \cdot40$).

**Discussion**

The current study had three aims. Firstly, to find a causal link between sustained attention and word production by using a manipulation known to tax sustained attention...
from the attentional literature. Event rate was manipulated in both a sustained attention task and picture naming task to see if this would result in corresponding effects. A second aim was to find out whether differences between a fast and slow event rate are due to differences in attention depletion, adjustments in focusing of attention, or whether it reflects a speed-accuracy trade-off. Thirdly, to show sustained attention ability correlates with language production, even in a simple naming paradigm.

The first aim was to find out if taxing sustained attention in a picture naming task would result in impaired performance. An event rate manipulation was used in both a picture naming task and in traditional sustained attention task. The prediction was that the manipulation known to tax sustained attention would result in similar effects for both tasks. Specifically, most previous studies have reported worse performance for fast event rates than slow event rates thought to be due to attention depletion, which should be reflected in lower hit rates, more false alarms, and a larger vigilance decrement. Here the event rate manipulation resulted in highly comparable outcomes in the sustained attention task and picture naming task, but not in the predicted direction. In both tasks participants were faster to respond when they were presented in rapid succession as compared to a slower presentation rate. Moreover, in both tasks, there was a vigilance decrement in RTs for both event rates, but there was no difference in the magnitude of the decrement between the two event rates. The only difference between the two tasks was in accuracy levels, such that hit rates were lower for the fast event rate of the DDT than for the slow rate, whereas there were hardly any errors made when naming pictures for either rate.

The fact that hit rates were lower for the fast event rate than the slow event rate in the DDT seems to provide evidence for the attention depletion hypothesis. However, RTs were also faster for the fast event rate, which could point to a speed-accuracy trade-off instead. To prove that fast event rates are more taxing one needs to show a larger decrement in hit rates and/or RTs in the fast condition as compared to the slow event rate. There was a vigilance decrement for hit rates in both conditions, but no interaction. Similarly, a decrement in the RTs was present, but again no interaction with event rate, suggesting the decrement was similar for both event rates. Some previous studies found a larger number of false alarms for the slow event rate, which would argue against a speed-accuracy trade-off, but the current sustained attention task did not show this effect. The current result seems to speak in favor of a speed-accuracy trade-off interpretation. Whether speed or accuracy is prioritized depends on both the perceptual input but also on environmental constraints and internal goals (Heitz, 2014). In the fast event rate, there is less time for evidence to accumulate to decide whether the stimulus is a target or a non-target. As a result, participants are faster to respond but also make more errors.

It should be noted that for the DDT, the number of events per minute in the fast event rate was well above the cut-off of 48 events per minute as suggested by Lanzetta et al. (1987), namely 100 events per minute. Therefore, one cannot argue that the fast event rate condition was actually too slow. All in all, a simple trade-off explanation for the DDT cannot be refuted. These findings warrant caution for interpreting fast event rates as more taxing on sustained attention than slow rates, as done by the attention depletion account. Experiments need to report both hit rates and RTs and show that a decrement for one measure does not go hand in hand with an improvement for the other. Only reporting errors, as has been done predominantly in past research on sustained attention (Ballard, 2001; Coulh et al., 1996; Jerison & Pickett, 1964; Lanzetta et al., 1987; Parasuraman, 1979; Parasuraman & Giambra, 1991), will not contest the speed-accuracy trade-off.

An attention depletion account also does not seem to hold for the event rate manipulation in the picture naming tasks. Accuracy levels for both event rates was near ceiling at 99% but responding was faster in the fast event rate condition. In contrast, the naming results support the proposal that the differences between fast and slow event rates are due to an adjustment in focusing of attention, in line with De Jong et al. (1999). They found faster RTs without a decline in accuracy for the fast event rate in a spatial Stroop task as compared to a slow rate. They argued that the fast event rate helped to keep focus on the task at hand, whereas the slow event rate gave rise to more fluctuations in attentional state, allowing word reading to interfere with responding to the location of the word. It could thus be the case that certain tasks benefit from a fast presentation rate, perhaps tasks where a response has to be given on each trial (as in the picture naming task) instead of only on a subset of trials as in the traditional sustained attention tasks (i.e., the DDT). Another possibility is that tasks that use more complicated and/or linguistic stimuli benefit from a fast event rate whereas simple non-linguistic tasks do not. So far, only De Jong et al. and the present study have used linguistic stimuli with complicated tasks (word reading and picture naming respectively). All previous studies testing sustained attention use tasks like the DDT, where stimuli are easier to decode (number or letter) and responses are easier to compute (button press). A benefit from focused attention due to a fast event rate might only be revealed for difficult tasks.

The lack of errors in both event rates could also indicate that the event rate manipulation in the picture naming task was less effective than in the sustained attention task. Accuracy was not analyzed statistically as error rate was only just above 1% but there is a hint that more errors and hesitations were made in the fast than slow event rate. This would follow the pattern of the DDT and again point to a speed-accuracy trade-off. It is also possible that the event rate manipulation did not work at all and therefore cannot be used as evidence against the attention depletion account. The lack of naming errors in both event rates could indicate that both event rates were relatively easy and that the fast event rate was actually not fast enough to strongly tax attention. The ISI was very short for a picture naming task, namely 500 ms. However, the picture stayed on the screen for one second. Picture duration was chosen
at one second to give people enough time to identify the object, plan the name, and complete the speech output before the next trial started. The total trial length of 1.5 seconds may have been too long to tax sustained attention to a larger degree than the 3 second trial length in the slow event rate. With 40 events per minute, the fast event rate fell just below the cut-off of 48 events per minute for simple tasks as suggested by Lanzetta and colleagues (1987). It could be that the naming task is indeed such a simple task. Participants are familiarized with all the items, and repeat all items thirty times. The highly repetitive nature could have made the task too simple and participants had few problems naming pictures, even in the fast event rate.

Previous research has used similar ISIs to differentiate between fast and slow conditions as used in the present study (Smallwood et al., 2004). Not finding a larger vigilance decrement in the fast event rates could be due to other task parameters, such as task duration. Each experiment lasted approximately 30 minutes, with blocks around 7 minutes. It could be that the blocks did not last long enough to thoroughly tax sustained attention. Another possibility is that the alternation of fast and slow blocks caused participants to recharge, and as such the second block was not more difficult than the first block. It could be the case that within blocks, the performance decrement was larger for the fast event rates. Post-hoc analyses, adding a factor to the model dividing blocks into two, did not provide evidence for this idea. For the DDT, the interaction between event rate and this new factor did not reach significance for either hit rates ($\beta = 0.03, SE = 0.24, z = 0.14, p = 0.89$) or RTs ($\beta = -0.00, SE = 0.01, t = -0.20$). There was no interaction effect for picture naming either ($\beta = 0.00, SE = 0.00, t = 0.38$).

The third aim was to link this study to previous research on sustained attention and its role in language production by correlating individuals’ sustained attention ability with picture naming performance. In the previous studies, the correlation was strongest for production latencies when naming occurred in a relatively difficult situation, i.e. a dual-task experiment (Jongman, Meyer, & Roelofs, 2015; Jongman, Roelofs, & Meyer, 2015). The correlation with simple picture naming was significant, but weaker (Jongman, Roelofs, & Meyer, 2015). Here, by increasing the need to sustain attention within a simple picture naming paradigm instead of using a more demanding naming paradigm, sustained attention ability correlated with single word production. Both event rates were faster than the event rate used in the only study testing the role of sustained attention in simple picture naming, and participants named many more pictures in the current experiment (900 versus 240). This very likely made the current task more difficult than the simple picture naming task in the study by Jongman, Roelofs and Meyer. As in earlier research, the sustained attention task was found to correlate with the right tail end of naming latency distributions. However, the sustained attention task also correlated with the normal portion, suggesting the picture naming task was so demanding that individual differences became evident for the whole RT distribution and not just a subset of responses. Thus, individuals with worse sustained attention were consistently slower when naming pictures ($\mu$) and showed a larger amount of very slow responses ($\tau$) as compared to individuals with better sustained attention.

It is a possibility that the relationship between the mean RT on the DDT and the $\mu$ parameter for picture naming is not due to high task demands. The correlation could instead be due to a general reduction in speed of processing in certain individuals, instead of being due to a specific decrease in their sustained attention ability. This should be tested by adding a third task that measures general speed of processing. However, a similar argument cannot be made for the correlation between the mean RT on the DDT and the $\tau$ parameter of picture naming. General speed of processing should affect only the mean response time, not the size of the right tail of an RT distribution. Instead the correlation fits with the idea that $\tau$ reflects lapses of attention as previously proposed by Unsworth et al. (2010). Moreover, the fact that two dissociable measures of sustained attention, RT and hit rate, both show a correlation with the $\tau$ parameter of picture naming also supports the idea that picture naming is indeed tapping into sustained attention. Hit rate, the proportion of correctly detected targets, was lower than we found previously when using a DDT with an intermediate event rate. This allowed for testing the correlation between hit rates and picture naming latencies, and a significant relationship was found for the $\tau$ parameter. It shows that hit rates can quantify sustained attention ability if there are strong individual differences, and that not only RTs on the DDT correlate with picture naming latencies but also accuracy levels. This provides stronger evidence that picture naming calls upon sustained attention.

In conclusion, the main aim was to show that simple picture naming requires sustained attention by manipulating a task parameter that is often used in sustained attention tasks. The task parameter, event rate, did not lead to the predicted result of fast event rate being more taxing than the slow event rate. However, performance was similar across the sustained attention task and picture naming, suggesting similar mechanisms are at play in both tasks. A manipulation that undoubtedly taxes sustained attention should be used in a picture naming task to show a definitive causal link between sustained attention and production. The correlational results do suggest that this link is present as sustained attention ability, measured not only by RTs but also hit rates, correlated with the abnormally slow responses when naming pictures, even in a simple picture naming paradigm. In sentence or discourse production the effect of poor sustained attention could be much more pronounced, causing slow speech, many hesitations and possibly even errors. Future research is necessary to test the role of sustained attention in more natural language production.

**Additional Files**
All the stimuli, presentation materials, participant data, and analysis files can be found on this paper’s project page on figshare: https://doi.org/10.6084/m9.figshare.4685038.v1.
Object names picture naming tasks. Target names of pictures, with English translation. DOI: https://doi.org/10.1525/collabra.84.s1

Additional analyses. Performance in the slow event rate conditions compared to performance on the equal amount of trials in the fast event rate conditions. DOI: https://doi.org/10.1525/collabra.84.s2

Acknowledgements
I thank Gurupriya Ramanathan and Hanna de Groot for help with setting up the experiment, testing the participants, and annotating part of the data, and Antje Meyer, Ardi Roelofs, and Brónagh McCoy for valuable comments on an earlier draft of this paper.

Competing Interests
The author has no competing interests to declare.

References


