
The Processing of Color Words in Sentence Comprehension

Journal:	<i>Quarterly Journal of Experimental Psychology</i>
Manuscript ID	QJE-STD-25-221.R1
Manuscript Type:	Standard Article
Date Submitted by the Author:	16-Nov-2025
Complete List of Authors:	Buchner, Emily; University of Würzburg, Department of Psychology IV Richter, Tobias; University of Würzburg, Department of Psychology IV Lenhard, Wolfgang; University of Würzburg, Department of Psychology IV
Keywords:	color representation, perceptual simulations, reading, sentence comprehension

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14 **The Processing of Color Words in Sentence Comprehension**
15

16
17
18 Emily L. Buchner, Tobias Richter, and Wolfgang Lenhard
19

20 Department of Psychology IV, University of Würzburg
21
22
23
24
25
26
27
28
29
30
31

32 **Author Note**
33

34 Emily L. Buchner, <https://orcid.org/0009-0002-6830-3713>
35

36 Tobias Richter, <https://orcid.org/0000-0002-0467-9044>
37

38 Wolfgang Lenhard, <https://orcid.org/0000-0002-8184-6889>
39
40
41
42

43 The experimental text material, data files, and R scripts for the analyses are available
44 at https://osf.io/7h2py/?view_only=f64033c9cd6a4d8abd2c8c33e6124ca2. All experiments
45
46 were preregistered.
47
48

49 Address for correspondence: Emily Buchner, University of Würzburg, Department of
50
51 Psychology IV, 97070 Würzburg, Germany. Email: emily.buchner@uni-wuerzburg.de
52
53
54
55
56
57
58
59
60

Abstract

This study investigated how color is represented in language comprehension. Some theories suggest that perceptual simulations—activations of sensory features such as shape, size, or color—routinely support understanding during reading. One line of evidence is the “mismatch effect” in sentence-picture verification tasks: responses are slower when pictures mismatch perceptual details described in the sentence. Across three preregistered experiments ($N = 222$), we tested whether this mismatch effect occurs with explicit color words (Experiment 1) and whether background colors presented concurrently interfere with the mental simulation of color to test the functionality of mental simulations in language comprehension (Experiments 2 and 3). As expected, participants responded faster when pictures matched the sentence’s color across the three experiments. When conflicting background colors were introduced during sentence presentation, the mismatch effect remained unaffected. This pattern of findings suggests that colors are routinely activated through color words during comprehension but the functional role of perceptual simulations of color for comprehension remains unclear.

Keywords: color representation; perceptual simulations; reading; sentence comprehension

The Processing of Color Words in Sentence Comprehension

Consider the following example: “Timothy opens the blue lid of the lunchbox and sees a green layer where he had placed a clementine a fortnight ago.” The sentence contains both colors that are explicitly named (blue, green) and a color that is implicitly associated (clementine, typically orange). Understanding these color references enables readers to represent what the lunchbox looks like and to infer that the clementine has grown moldy.

Some theories of language comprehension assume that linguistic information activates multimodal networks that store experiential properties, resulting in a multimodal simulation (Barsalou, 1999, 2008; Zwaan, 2004). Applied to the comprehension of color information, a color word should lead to a simulation of the described color. One way to examine the assumption that perceptual representations are activated during language comprehension is the sentence-picture verification paradigm (Stanfield & Zwaan, 2001). In this paradigm, pictures are presented after sentences, and the participants’ task is to judge whether the object depicted in the picture appeared in the sentence. By varying perceptual properties such as the orientation (Stanfield & Zwaan, 2001) or the color of the picture (Hoeben Mannaert et al., 2017) studies have shown that the verification response is faster for matching than for mismatching perceptual properties, which has been interpreted as meaning that perceptual properties are routinely activated during sentence comprehension (in line with Barsalou, 1999). Furthermore, neuroimaging studies have demonstrated that the brain areas recruited for processing perceptual properties overlap with brain areas recruited for processing words referring to these properties (e.g., Kosslyn et al., 2001; for colors: Simmons et al., 2007; Van Dam et al., 2012). Nevertheless, it remains unclear whether perceptual representations are functional for comprehension or merely a by-product of the core comprehension process.

1
2
3
4
5 The present experiments examined the role of perceptual simulations of color in
6
7 sentence comprehension with the two primary objectives of investigating whether color
8
9 information is routinely activated when explicit color words are provided and investigating
10
11 the functional role of perceptual simulations in the comprehension of color words. In the
12
13 following sections, we briefly review the current state of the literature and previous
14
15 methodological approaches to investigating color representations in language comprehension.
16
17 Based on this review, we derived specific research questions, which we addressed in three
18
19 experiments employing sentence-picture verification tasks. First, we examined whether the
20
21 well-established mismatch effect also applies to color words (see Hoeben Mannaert et al.,
22
23 2017, 2021; Zwaan & Pecher, 2012, for implicit color information). Second, we introduced
24
25 conflicting visual color information concurrently with sentence processing to gain insights
26
27 into the mechanisms underlying mental simulations. Comprehension effects were assessed in
28
29 terms of response latencies and accuracy.
30
31
32
33

34 **Studies Using the Sentence-Picture Verification Paradigm to Study the Role of Color in** 35 36 **Comprehension** 37

38
39 According to the theory of perceptual symbol systems (Barsalou, 1999, 2008),
40
41 perceptual properties of an object are routinely represented mentally when language is
42
43 comprehended. An established method for investigating the representation format of
44
45 language is the sentence-picture verification task (cf. Stanfield and Zwaan, 2001):
46
47 Participants read a sentence including an object that implies certain perceptual properties.
48
49 They then see a picture that either matches or mismatches the perceptual properties of the
50
51 object. Their task is to judge whether the depicted object was mentioned in the sentence.
52
53 Typically, this results in the so-called match advantage (e.g., Hoeben Mannaert et al., 2017),
54
55 also known as the mismatch effect (e.g., Zwaan et al., 2002), which means that responses to
56
57 the pictures that represent incongruent perceptual properties are slower than to the pictures
58
59
60

1
2
3
4
5 that represent congruent perceptual properties. The effect is interpreted as indicating a routine
6
7 activation of perceptual properties and as support for the assumption that comprehension
8
9 processes are accompanied by perceptual simulations. The mismatch effect has been
10
11 demonstrated across multiple perceptual dimensions (for a collection of experiments, see
12
13 Zwaan & Pecher, 2012) such as object shape (Engelen et al., 2011; de Koning et al., 2017a;
14
15 Zwaan et al., 2002) or size implied in sentences (de Koning et al., 2017a, 2017b; Koster et al.,
16
17 2019).

21 For the perceptual dimension of color, the results have been less conclusive. Color is
22
23 considered a perceptual dimension defined unambiguously by visual properties. That is, its
24
25 representation is assumed to not be compensated by non-visual modalities (Ostarek &
26
27 Bottini, 2021). The investigation of this perceptual dimension could therefore be of particular
28
29 interest for uncovering simulation mechanisms. In the past, however, inconsistent results
30
31 were obtained when using sentence-picture verification tasks with objects that imply certain
32
33 colors (color-diagnostic objects such as coal, tomatoes, or leaves). Contrary to the mismatch
34
35 effect found for other perceptual dimensions, Connell (2007) found faster responses to
36
37 incongruent combinations compared to congruent combinations. However, in a direct
38
39 replication of Connell's study (using the same stimulus material) Zwaan and Pecher (2012)
40
41 found a mismatch effect, which is consistent with results usually obtained for other
42
43 perceptual dimensions. As one possible explanation for these divergent results, Liu et al.
44
45 (2024) discussed whether individual differences exist in the representation of implicit color
46
47 information (e.g., green vs. orange for a leaf) because of differences in prior experiences.
48
49 Another possibility is that color is processed and represented in a different way, as some
50
51 studies showed that color is represented less stable or requiring more attentional resources
52
53 than the other perceptual dimensions during comprehension (for a discussion, see Connell,
54
55 2007). In a conceptual replication based on a new stimulus set (again with implicit color
56
57
58
59
60

1
2
3
4
5 information), the mismatch effect for color was replicated again (Hoeben Mannaert et al.,
6
7 2017, Experiment 1). Along the same lines, a comparison of different perceptual dimensions
8
9 (specifically shape, size, color, orientation) also yielded a greater mismatch effect for color
10
11 compared to the other dimensions investigated (de Koning et al., 2017a). Furthermore, color
12
13 representations were found to be rich in perceptual detail as the mismatch effect disappeared
14
15 when the saturation of the picture color was reduced (Hoeben Mannaert et al., 2017,
16
17 Experiments 2a and 2b).
18
19

20
21 To summarize, the majority of studies support the perspective that perceptual
22
23 simulations of colors are routinely activated—at least for implicit color information.
24
25 However, individual differences in people’s past experiences might be particularly relevant
26
27 for colors, which could explain the contradictory results (cf. Connell, 2007; Zwaan & Pecher,
28
29 2012). Additional insights into the processing of verbal color information could be gained by
30
31 investigating explicit color words for which individual differences in past experiences should
32
33 be less relevant. To our knowledge, the only study using explicit color words in sentence-
34
35 picture verification tasks was conducted by Hoeben Mannaert et al. (2021) who investigated
36
37 the activation of color across multiple sentences. They varied sentences with explicit vs. no
38
39 color references and with pictures that showed a color vs. a grey-scale variant of the focal
40
41 object. Instead of employing the classical mismatch effect, they investigated a color
42
43 advantage, that is, whether participants responded faster to colored pictures when a color
44
45 reference was included in the sentence than to pictures with greyscale, whereas no difference
46
47 in response latencies was expected when the sentence contained no color reference. The color
48
49 advantage appeared as expected, even across multiple sentences but only when the focus was
50
51 on the target object.
52
53
54
55
56
57
58
59
60

Overlaps in the Visual Processing of Colors and the Comprehension of Color Words

According to Barsalou's theory (1999, 2008), the activated brain areas for the visual processing and the word-based processing of perceptual properties should overlap. In support of this assumption, activations in color processing areas were found for the reception of explicit color words (Simmons et al., 2007) and for implicitly associated color properties (Van Dam et al., 2012), which can be clearly distinguished from activations elicited by other perceptual dimensions (for a comparison with form words, del Prado Martín et al., 2006; Pulvermüller & Hauk, 2006). Studies have shown that words activate brain areas associated with their perceptual word meaning for other dimensions. For example, when the meaning of words is associated with visual properties, the visual cortex is activated during the processing of the words (e.g., Kosslyn et al., 2001; Mak & Willems, 2021; Pulvermüller, 1999). Behavioral experiments extended these correlational findings by demonstrating that the presentation of competing visual stimuli can interfere with the recall of visual knowledge associated with words (Edmiston & Lupyan, 2017) and impair the speed and accuracy of semantic judgements on words for which the content is associated with visual characteristics, with visual interference being more effective the more subjective visual experience the participants had with the object (Davis et al., 2020). This methodological approach has also been applied to studying the functionality of perceptual simulations in language comprehension, most explicitly for the perceptual dimension of shape (Ostarek et al., 2019) and dynamic directions (Buchner et al., 2025), but to our knowledge not yet for the dimension of color.

However, behavioral studies based on other paradigms that leverage interference also suggest an interplay between the processing of colors and color words. One example is the Stroop effect (Stroop, 1935; for an overview see MacLeod, 1991), which is characterized by an increase in errors and processing time when naming the ink color of a presented word that

1
2
3
4
5 is incongruent with the color word. However, this design addresses attentional processes, the
6
7 automaticity of reading processes compared to color naming, and the parallel processing of
8
9 relevant and irrelevant stimuli (for a discussion see MacLeod, 1991) rather than the mental
10
11 representation elicited by the color words. In the context of sentence comprehension, a
12
13 modified semantic Stroop task was used by Connell and Lynott (2009) to investigate which
14
15 colors are mentally represented when a sentence only implies color information (e.g., a bear
16
17 at the North Pole indicating a bear with white fur). After participants read a sentence, the
18
19 object word appeared again written in colored ink. The naming of the ink color was slower
20
21 when the color mismatched the information from the previous sentence compared to when
22
23 naming matching colors. Thus, the processing of implied color information appeared to
24
25 facilitate color processing of visually congruent colors. A variety of Stroop-associated
26
27 interference tasks have been developed, some of which have been extended with priming
28
29 paradigms and expanded to more general picture-word interference (for an overview, see
30
31 Glaser, 1992). Word-color interference tasks in which color primes are followed by color
32
33 words to form congruent and incongruent trials (Cui et al., 2007; replication study by Azañón
34
35 et al., 2025) are particularly informative for the present experiments. These studies
36
37 demonstrated that trials with congruent color primes and color words lead to faster and more
38
39 accurate responses compared to incongruent trials (see Azañón et al., 2025). Furthermore,
40
41 Berndt et al. (2020) found that color-congruent primes (e.g., green) reduced the solution time
42
43 for anagrams (e.g., *cbemcuru* for *cucumber*), which the authors interpreted as evidence for an
44
45 important role of multimodal experiential traces in language comprehension.
46
47
48
49
50
51

52
53 These results are extended by the findings of Richter and Zwaan (2009) who found
54
55 bidirectional congruency effects in the processing of visually perceived colors and color
56
57 words. They combined a color discrimination task with a presentation of matching or
58
59 mismatching color words on which lexical decisions were made. Lexical decisions on color
60

1
2
3
4
5 words were found to be slower when the color word was incongruent with the color seen
6
7 before compared to congruent combinations. In addition, color discrimination judgments
8
9 were found to be slower when the colors were incongruent with a color word seen before
10
11 compared to congruent combinations. This effect persisted when the lexical decision task was
12
13 omitted, suggesting that color representations are activated routinely when color words are
14
15 read.
16
17

18
19 To summarize, neuro-imaging and behavioral evidence suggests that the simulation of
20
21 color properties based on words and the representation of colors based on perceptual
22
23 impressions are interconnected. However, the question of the extent to which perceptual
24
25 simulations are functional for comprehension remains unclear. Existing literature lacks
26
27 experimental approaches that inhibit the simulation of color and objectify the effects on
28
29 comprehension measures.
30
31

32 **Rationale of the Present Experiments**

33
34
35 The present experiments were designed to examine the representation of color in
36
37 sentence comprehension. We pursued two objectives. The first objective was to investigate
38
39 whether the mismatch effect, that is, incongruent perceptual information in a sentence-picture
40
41 verification task leads to impaired object recognition compared to congruent perceptual
42
43 information (Zwaan & Pecher, 2010), occurs for the perceptual dimension of color when
44
45 sentences contain explicit color words (Experiment 1). We expected that responses in a
46
47 sentence-picture verification task would be significantly slower (Hypothesis 1) and less
48
49 accurate (Hypothesis 2) when the color of the object shown (picture color) and the color
50
51 explicitly mentioned in the preceding sentence (color word) were incongruent compared to
52
53 when the picture color and the color word were congruent.
54
55
56

57
58 The mismatch effect may be interpreted as evidence for the activation of modal
59
60 representations in comprehension (e.g., Stanfield & Zwaan, 2001; Zwaan et al., 2002; Zwaan

1
2
3
4
5 & Pecher, 2012) and as an indicator for mental simulations including multiple perceptual
6
7 object properties (e.g., Hoeben Mannaert et al., 2017), but it cannot be interpreted as evidence
8
9 for a functional role of mental simulations for successful comprehension (Mahon &
10
11 Caramazza, 2008; for an overview, see Ostarek & Huettig, 2019). Therefore, the second
12
13 objective was to gain a better understanding of the functionality of perceptual simulations in
14
15 sentence comprehension. Neuroimaging studies indicate overlapping brain activation for
16
17 language comprehension and perceptual processing (e.g., Kosslyn et al., 2001; Mak &
18
19 Willems, 2021; Pulvermüller, 1999). This observation has also been made for color words
20
21 and color processing areas (Simmons et al., 2007) as well as for words with specific color
22
23 properties and color processing areas (Van Dam et al., 2012). These findings are
24
25 correlational, but some experimental evidence also shows that the retrieval of visual
26
27 information associated with the word meaning can be disrupted by visual stimuli (for single
28
29 words, see Davis et al., 2020; Edmiston & Lupyan, 2017; Ostarek & Huettig, 2017; Richter &
30
31 Zwaan, 2009). For color representations, these findings lead to the assumption that the
32
33 presentation of visual color information can disrupt the simulation of word-based color
34
35 information because both sources of information compete for the same cognitive resources.
36
37
38
39
40

41 In Experiment 2, we used the same sentence-picture verification task from
42
43 Experiment 1 but also included an experimental manipulation of color during comprehension.
44
45 Colored background screens during the sentence presentation (background color) were used
46
47 to visually display a congruent or incongruent color relative to the color word. If perceptual
48
49 simulations are indeed functional for comprehension, an incongruent background color
50
51 should hamper comprehension of the color words, leading to slower and less accurate
52
53 responses in the sentence-picture verification task compared to a congruent background color.
54
55 We expected slower and less accurate responses in the verification task when the background
56
57 color was incongruent with both the color word and the picture color (incongruent
58
59
60

Method

Participants

We conducted an a-priori power analysis for a one-tailed *t*-test as an approximation of the planned (Generalized) Linear Mixed Models, under the expectation of a small effect of Cohen's $d = 0.30$ (cf. Zwaan et al., 2002), a Type I error probability of $\alpha = .05$, and a power of $1 - \beta = .80$. Adding a margin of 20% for pre-specified exclusions and to acknowledge that (Generalized) Linear Mixed Models may require a larger sample, this resulted in a required sample size of $N = 84$ (according to G*Power, Faul et al., 2007). The following preregistered exclusion criteria were applied in all three experiments: First, participants were only eligible to take part in the experiment when they were able to speak and understand German on a native level and were able to fully distinguish colors. Second, participant data were excluded when less than 80% (cf. Hoeben Mannaert et al., 2017, 2021) of the sentence-picture verification task responses were correct. Third, we preregistered that when the average accuracy for individual sentence-picture combinations was below 80%, this combination was excluded for analyses (cf. Hoeben Mannaert et al., 2017). Fourth, we log-transformed response latency (ms) and only included data within a range of $M \pm 2 SD$ of the log-transformed response latency based on group means. The third criterion did not apply to any of the three experiments reported in the current study.

Four participants were excluded because of the second exclusion criterion. Eighty participants were included in the analysis (age: $M = 25.33$ years, $SD = 8.70$; 65 female, 14 male, 1 diverse). The majority of the participants were students ($n = 73$) of which 20 were psychology students, 34 were teacher students, and the remainder ($n = 19$) were enrolled in other study programs. The participants were offered course credit or monetary reward for participation in all experiments.

Materials

Sentences. We created 72 German sentences for the sentence-picture verification tasks (following Stanfield & Zwaan, 2001). Each sentence mentioned a focal object. None of the objects implied a specific color but were color-neutral (e.g., a ball, a jumper, an umbrella) and were complemented by a color word. The color words appeared either in the third or sixth position of the sentences, restricted by the grammar of German. We used four color words (yellow, red, blue, green), distributed equally across the two sentence structures. Each sentence consisted of seven words.

Pictures. Pictures corresponding to the sentences were created with the illustration software SketchBook (version 5.1.0). The pictures depicted objects schematically but in color. The colors used in the procedure represented pure instances of the RGB (Red, Green, Blue) color model for yellow (255, 255, 0), red (255, 0, 0), green (0, 255, 0), and blue (0, 0, 255), which were related to the four colors described in the sentences.

For experimental trials, the pictures represented the objects described in the corresponding sentences. Each picture existed in two color versions: congruent to the color word in the corresponding sentence and incongruent to the color word in the corresponding sentence. For filler trials, the pictures represented objects other than those described in the corresponding sentences. The colors illustrated in the filler pictures were always congruent with the color word in the corresponding sentence (to prevent strategic responding, cf. van Zuijlen et al., 2024).

Pilot Study. The stimuli were pilot-tested with an independent sample of 10 participants. The aim of the pilot study was to validate whether the objects visualized in the pictures were identifiable in combination with the experimental sentences. One participant responded “not at all” to the question “Was it clear to you what the pictures were supposed to represent?” and was thus excluded. Nine participants were included (age: $M = 22.44$ years,

1
2
3
4
5 $SD = 2.96$; 8 female, 1 male). We presented 84 sentence-picture combinations, divided into
6
7 56 experimental trials and 28 filler trials. Participants read a sentence followed by a picture
8
9 that depicted the congruent color. They were asked to judge whether the object visualized
10
11 matched the object described in the preceding sentence. Participants were instructed to press
12
13 “j” when the object matched the object described in the sentence and “f” when it did not
14
15 match. When less than 85% of the participants gave an unambiguous answer for the
16
17 experimental sentence-picture combinations, this combination was revised for the main
18
19 experiments. This procedure resulted in the revision of five sentence-picture combinations.
20
21 One additional object (“eyeshadow”) was revised because we received feedback from
22
23 participants that this word could cause gender-specific differences in response behavior.
24
25
26

27
28 We also asked for a general assessment of the sentences and pictures, using a 7-point
29
30 Likert scale (0 = *not at all*; 7 = *very good*). The quality of the pictures ($M = 6.56$, $SD = 0.72$)
31
32 and the comprehensibility of the sentences ($M = 6.11$, $SD = 0.93$) were rated as high. We also
33
34 asked whether the number of sentences and pictures was acceptable. Agreement was very
35
36 high ($M = 6.44$, $SD = 1.33$).
37
38

39
40 **Dependent Variables.** Dependent variables were response latency and response
41
42 accuracy in the sentence-picture verification tasks. Analyses of response latency were based
43
44 on log-transformed response times (in ms), and only correct responses were included.
45

46 **Procedure**

47
48 The experiment was conducted in the laboratory with the software Inquisit Lab
49
50 (version 6.6.1, Millisecond, 2021). Participants sat in front of a computer with an external
51
52 screen and a keyboard. Participants received the instructions in written form via the computer
53
54 screen, gave their informed consent, and provided demographic data. We conducted a brief
55
56 color diagnostic, which explained to the participant as a keyboard test because we tried not to
57
58 draw attention to color reception throughout the communication of the experiment to avoid
59
60

1
2
3
4
5 unintended effects. In this context, two pictures from the ensemble of the Ishihara test (for a
6
7 discussion and evaluation, see Birch, 1997) and two pictures showing letters appeared.

8
9 Participants were required to press the corresponding keys for the numbers and letters on the
10
11 screen, which ensured that all participants were able to discriminate colors beyond self-
12
13 assessment.
14

15
16 The instructions for the experimental task followed Hoeben Mannaert et al. (2021).
17
18 Participants should rate whether the object mentioned in the sentence was depicted in the
19
20 picture. However, they were instructed to disregard the color in their judgments. Responses
21
22 were to be given via key press (“j” for “true” or “f” for “false”) with the index fingers of the
23
24 left and the right hands. Furthermore, participants were instructed that each word was
25
26 presented separately and that they should press the space bar with their thumb when they
27
28 wanted to proceed to the next word. As in Zwaan et al. (2002), participants were told that
29
30 they should respond as quickly as possible because their response times would be measured.
31
32 After reading the written instructions, participants practiced the task on eight sentence-picture
33
34 combinations with feedback (as in Hoeben Mannaert et al., 2017; Stanfield & Zwaan, 2001)
35
36 and afterwards started with the main experiment. We reminded the participants to place their
37
38 fingers on the relevant keys of the keyboard.
39
40
41
42
43

44 In the main part of the experiment, 72 sentence-picture verification tasks were
45
46 presented (without feedback). Each trial appeared in the same sequence (see Figure 1). A left-
47
48 aligned fixation cross appeared on the screen for 1000 ms, followed by the first word (black
49
50 letters on a white screen). By pressing the space bar, participants moved to the next word.
51
52 When all words had been displayed, a fixation cross appeared in the center of the screen for
53
54 500 ms. Then a picture was shown in the center of the screen. Participants responded with “j”
55
56 for “true” or “f” for “false”. After the participants’ response, the picture disappeared, and
57
58 then a blank screen was shown for 500 ms before the next trial began.
59
60

1
2
3
4
5 In half of the experimental trials (24 trials), the picture color was congruent to the
6
7 color word (congruent condition). In the other 24 trials, the picture color was incongruent to
8
9 the color word (incongruent condition). Each of the four color words (yellow, red, blue,
10
11 green) occurred 12 times in the experimental trials. Each condition (congruent/incongruent)
12
13 appeared six times per color, with the color word appearing three times in third place and
14
15 three times in sixth place. All experimental trials (congruent and incongruent condition) were
16
17 to be answered with “j” for “true”.
18
19

20
21 In the 24 filler trials, the object depicted was different from the object described. Each
22
23 of the four color words occurred six times. All filler trials were to be answered with “f” for
24
25 “false”.
26
27

28 For the experimental trials, two lists were created to ensure that each sentence
29
30 occurred equally in combination with a congruent picture color and an incongruent picture
31
32 color across participants. Within each list, the order of sentence presentation was randomized.
33
34 In both lists, all fillers were additionally presented in a randomized manner. Participants were
35
36 randomly assigned to one of the two lists.
37
38

39 After the experiment, the participants were presented with sentence verification tasks,
40
41 focusing on microstructural reading processes (ELVES; Richter & van Holt, 2005), a mental
42
43 comparison task that focused on mental imagery abilities (MCT; Suggate, 2024), and a
44
45 questionnaire on mental imagery abilities (Psi-Q; Andrade et al., 2014). The data are included
46
47 in the data files on OSF. These tasks were not relevant to the research questions of the current
48
49 study. Thus, no results are reported for these tests. The entire experiment took about 30 min.
50
51

52 ***Design***

53
54 The experiment was based on a one-factorial within-subject design with color match
55
56 vs. mismatch as the independent variable (congruent vs. incongruent colors of object picture
57
58 and color word).
59
60

Data and materials availability

The full data, the R code for the model results as well as the experimental text material are available at the repository of the Open Science Framework (https://osf.io/7h2py/?view_only=f64033c9cd6a4d8abd2c8c33e6124ca2). The full materials are available from the authors upon request.

Results and Discussion

The reported analyses were conducted in R (Version 4.4.1.). All hypotheses were tested with a Type I error probability of $\alpha = .05$. Linear Mixed Models (LMM) were estimated for response latency as a dependent variable (Baayen et al., 2008) and Generalized Linear Mixed Models (GLMM) were estimated for response accuracy as a dependent variable (Jaeger, 2008). The models included the main effect of incongruent vs. congruent condition (dummy coded 1 vs. 0). Random effects of participants and sentences were included (random intercepts). We used the R package lme4 to estimate the models (Bates et al., 2015). Follow-up analyses were conducted with the R package emmeans (Lenth, 2025).

Results revealed the mismatch effect predicted by Hypotheses 1 and 2 (Figure 2). Responses in the incongruent condition were significantly slower ($M = 6.580$, $SE = 0.032$) than in the congruent condition ($M = 6.484$, $SE = 0.032$), $\beta = 0.10$, $SE = 0.02$, $t(93) = 4.37$, $p < .001$ (one-tailed), $d = 0.37$. Responses in the incongruent condition were also slightly but significantly less accurate ($M = 0.985$, $SE = 0.004$) compared to the congruent condition ($M = 0.994$, $SE = 0.002$), $\beta = -0.92$, $SE = 0.32$, $z = -2.91$, $p = .002$ (one-tailed), $OR = 0.40$, 95% CI [0.21, 0.74].

Our results provide support for a routine, modality-based representation of color information in sentence comprehension when the color of the focal object is explicitly named in the sentence. The results correspond to the prototypical pattern of results in sentence-

1
2
3
4
5 picture verification tasks by revealing a mismatch effect for color (consistent with Hoeben
6
7 Mannaert et al., 2017; Zwaan & Pecher, 2012). However, this finding cannot be interpreted
8
9 as evidence for a functional role of color representations during comprehension. Therefore, in
10
11 Experiment 2, we provided competing color information during the reading process, which
12
13 offers the opportunity of gaining causal insights into the functionality of perceptual
14
15 simulations.
16
17

18 19 **Experiment 2**

20
21 The aim of Experiment 2 was to investigate whether the mismatch effect for colors
22
23 would be moderated by competing visually presented colors (background color) during the
24
25 reading process (for the preregistration see <https://aspredicted.org/649x-s5qk.pdf>). Based on
26
27 neuroimaging findings suggesting that perceptual impressions and the comprehension of
28
29 words that refer to the same perceptual dimension draw on the same cognitive resources
30
31 (Barsalou, 1999; for color, see Simmons et al., 2007; Van Dam et al., 2012), visually
32
33 presented conflicting colors should impair the comprehension of verbally described colors.
34
35 Findings from behavioral interference paradigms also support this assumption. These include
36
37 experiments with word-color interference tasks (for an overview, see Glaser, 1992). In this
38
39 context, bidirectional congruence effects were found for visually presented colors and color
40
41 words (Richter & Zwaan, 2009). Based on these findings we expected effects of conflicting
42
43 color information on comprehension to be reflected in an altered response pattern in the
44
45 sentence-picture verification task.
46
47
48
49

50
51 Specifically, we expected that when the picture color is congruent with the color word
52
53 but incongruent with the background color (incongruent background color condition),
54
55 participants would respond significantly slower (Hypothesis 1a) and less accurately
56
57 (Hypothesis 1b) to a sentence-picture verification task than when all described and depicted
58
59 colors were congruent (fully congruent condition). We also expected that when the picture
60

1
2
3
4
5 color is congruent with the background color but incongruent with the color word
6
7 (incongruent color word condition), participants would respond significantly slower
8
9 (Hypothesis 2a) and less accurately (Hypothesis 2b) in a sentence-picture verification task
10
11 than when all described and depicted colors are congruent (fully congruent condition). In the
12
13 incongruent color word condition, the background color and the picture color are visually
14
15 congruent, which is not the case in the incongruent background color condition. Therefore,
16
17 the background color might prime the congruent picture color, facilitating “yes” responses in
18
19 the sentence-picture verification task. Against this background, we assumed that responses in
20
21 the incongruent background color condition, which included visual color incongruency,
22
23 would be slower (Hypothesis 3a) and less accurate (Hypothesis 3b) than in the incongruent
24
25 color word condition, which included color congruency.
26
27
28

30 **Method**

32 *Participants*

34
35 We followed the power analysis for Experiment 1. Eighty-four data sets were
36
37 collected. One data set was excluded immediately after collection because the participant was
38
39 taking part for the second time. Eleven participants were excluded because they answered
40
41 less than 80% of the sentence-picture verification tasks correctly (according to our second
42
43 exclusion criterion), which led to a final sample size of 72 participants (55 female, 15 male, 2
44
45 diverse), with an average age of 26.99 years ($SD = 10.91$). The majority of the participants
46
47 were students ($n = 64$) of which 13 were psychology students, 23 were teacher students, and
48
49 the remaining students ($n = 28$) were enrolled in other study programs.
50
51

53 *Material*

55 The same material as in Experiment 1 was used.
56
57
58
59
60

Procedure

The procedure was almost identical to the first experiment. The only difference was in the presentation of the sentences (see Figure 3). The experimental manipulation appeared in the form of background colors. Within a sentence, the same background color was presented for each of the seven words. The background colors represented pure instances of the RGB color model for yellow (255, 255, 0), red (255, 0, 0), green (0, 255, 0), and blue (0, 0, 255), analogous to the picture colors. The background colors were either congruent or incongruent with the color word and the picture color. This manipulation resulted in three conditions with 16 trials each:

- Fully congruent condition: The color word, the picture color and the background color were congruent.
- Incongruent background color condition: The color word and the picture color were congruent, but the background color was an incongruent color.
- Incongruent color word condition: The picture color and the background color were congruent, but the color word was an incongruent color.

As in Experiment 1, 24 fillers were presented in which the object depicted differed from the object described. To keep the conditions parallel, we also displayed background colors for filler trials. In all filler trials, the background color matched the color word as well as the picture color. Again, all experimental trials were to be answered with “j” and all filler trials with “f”.

Each participant received 72 sentence-picture verification tasks. All sentences appeared once per participant. However, we ensured across all participants that each sentence appeared equally often in all three conditions. Therefore, we created three lists in which each sentence was assigned to another condition. Again, each of the four color words (yellow, red,

1
2
3
4
5 blue, green) occurred 12 times per list. Each condition (congruent/incongruent background
6
7 color/incongruent color word) appeared four times per color, with the color word appearing
8
9 twice in third place and twice in sixth place. The 24 fillers from Experiment 1 were also
10
11 presented and were identical for all participants. Participants were randomly assigned to one
12
13 of the three lists. The sentences were presented in random order.
14
15

16 ***Design***

17
18 The experiment was based on a one-factorial within-subject design with the
19
20 independent variable color congruency (fully congruent condition vs. incongruent
21
22 background color condition vs. incongruent color word condition).
23
24

25 **Results and Discussion**

26
27 We estimated an LMM for response latency and a GLMM for response accuracy
28
29 (random intercept models with random effects of participants and sentences) with condition
30
31 as a fixed factor (dummy-coded, with the fully congruent condition as the reference category
32
33 coded 0). Pairwise comparisons between conditions were obtained using Holm-corrected
34
35 contrasts (via the emmeans package). The means and standard errors for response latency and
36
37 response accuracy by conditions are shown in Figure 4. Please note that this procedure (valid
38
39 for Experiment 2 and 3) deviates slightly from the preregistrations, which describe the
40
41 estimation of three separate (G)LMMs. The preregistered procedure yielded the same pattern
42
43 of results, which are available in the Online Supplement
44
45
46
47
48
49 (https://osf.io/7h2py/?view_only=f64033c9cd6a4d8abd2c8c33e6124ca2).

50
51 For response latency, the comparison of the incongruent background color condition
52
53 ($M = 6.436$, $SE = 0.028$) with the fully congruent condition was not significant ($M = 6.412$,
54
55 $SE = 0.028$), $\beta = 0.02$, $SE = 0.02$, $t(138) = 1.15$, $p = .126$ (one-tailed). For response accuracy,
56
57 the comparison of the incongruent background color condition ($M = .995$, $SE = 0.002$) with
58
59
60

1
2
3
4
5 the fully congruent condition ($M = .996$, $SE = 0.002$), $\beta = -0.15$, $SE = 0.50$, $z = -0.30$, $p = .383$
6
7 (one-tailed) was also not significant. Instead, participants made hardly any errors in both
8
9 conditions. Thus, Hypotheses 1a and 1b received no support.
10

11
12 In a comparison of the incongruent color word condition and the fully congruent
13
14 condition (Hypotheses 2a and 2b), responses in the incongruent color word condition were
15
16 significantly slower ($M = 6.534$, $SE = 0.028$) than in the fully congruent condition ($M =$
17
18 6.412 , $SE = 0.028$), $\beta = 0.12$, $SE = 0.02$, $t(140) = 5.78$, $p < .001$ (one-tailed), $d = 0.51$.

19
20 Responses in the incongruent color word condition were also significantly less accurate (M
21
22 $= .981$, $SE = 0.006$) compared to the fully congruent condition ($M = .996$, $SE = 0.002$), $\beta = -$
23
24 1.58 , $SE = 0.41$, $z = -3.900$, $p < .001$ (one-tailed), $OR = 0.21$, 95% CI [0.08, 0.54]. This
25
26 pattern of results supports Hypotheses 2a and 2b.
27
28
29

30
31 In a comparison of the incongruent background color condition and the incongruent
32
33 color word condition (Hypotheses 3a and 3b), responses in the incongruent background color
34
35 condition were significantly faster ($M = 6.436$, $SE = 0.028$) than in the incongruent color
36
37 word condition ($M = 6.534$, $SE = 0.028$), $\beta = -0.10$, $SE = 0.02$, $t(139) = -4.63$, $p < .001$ (two-
38
39 tailed), $d = 0.41$. Responses in the incongruent background color condition were also
40
41 significantly more accurate ($M = .995$, $SE = 0.002$) compared to the incongruent color word
42
43 condition ($M = .981$, $SE = 0.006$), $\beta = 1.44$, $SE = 0.39$, $z = 3.71$, $p < .001$ (two-tailed), $OR =$
44
45 4.20 , 95% CI [1.67, 10.61]. Both effects run counter to the effects predicted in Hypotheses 3a
46
47 and 3b.
48
49

50
51 These unexpected results regarding Hypotheses 1 and 3 suggest that on the one hand,
52
53 congruent background color and object color do not override the interference introduced by
54
55 the incongruent color word in relation to the picture color. On the other hand, we found no
56
57 evidence that an incongruent background color hampers responses in the sentence-picture
58
59
60

1
2
3
4
5 verification task. This pattern of results suggests that the perceptual simulation of color
6
7 information is less influenced by concurrently presented background information. Instead,
8
9 color words that are incongruent with the picture color interfere with the sentence-picture
10
11 verification task irrespective of background color. To examine the reliability of this finding,
12
13 we conducted a direct replication study.
14
15

16 **Experiment 3**

17
18 Experiment 3 was a direct replication of Experiment 2 to follow-up on the unexpected
19
20 findings (preregistered at <https://aspredicted.org/qdpgk-n3kh.pdf>). As for Experiment 2, we
21
22 expected responses in the incongruent color word condition to be slower (Hypothesis 1a) and
23
24 less accurate (Hypothesis 1b) compared to the fully congruent condition. The results of
25
26 Experiment 2 furthermore suggested that an incongruent background color during sentence
27
28 presentation does not affect responses in the sentence-picture verification task. We therefore
29
30 expected responses in the incongruent color word condition to be slower (Hypothesis 2a) and
31
32 less accurate (Hypothesis 2b) compared to the incongruent background color condition.
33
34
35

36 **Method**

37 ***Participants***

38
39 We followed the power analysis for Experiment 1 and 2. Eighty-four participants took
40
41 part in the experiment. Fourteen participants were excluded because they answered less than
42
43 80% of the sentence-picture verification tasks correctly (according to our second exclusion
44
45 criterion). Data from 70 participants were included in the analysis (age: $M = 21.71$ years, SD
46
47 $= 3.73$; 58 female, 10 male, 2 diverse). All but one of the participants were students of which
48
49 20 were psychology students, 42 were teacher students, and the remaining students ($n = 7$)
50
51 were enrolled in other study programs.
52
53
54
55
56

57 ***Material***

58
59 The same materials were used as in Experiment 2.
60

Procedure

Experiment 3 followed the same procedure as Experiment 2.

Design

Experiment 3 was based on the same design as Experiment 2.

Results and Discussion

Again, we estimated an LMM for response latency and a GLMM for response accuracy (random intercept models with random effects of participants and sentences) with condition as a fixed factor (dummy-coded, with the fully congruent condition as the reference category coded with 0). Pairwise comparisons between conditions were obtained using Holm-corrected contrasts (via the emmeans package). The means and standard errors for response latency and response accuracy are shown in Figure 5.

The findings of Experiment 2 were replicated. Responses in the incongruent color word condition were significantly slower ($M = 6.634$, $SE = 0.028$) compared to the fully congruent condition ($M = 6.459$, $SE = 0.028$), $\beta = 0.17$, $SE = 0.02$, $t(140) = 8.34$, $p < .001$ (one-tailed), $d = 0.67$. Responses in the incongruent color word condition were also significantly less accurate ($M = .971$, $SE = 0.008$) compared to the fully congruent condition ($M = .995$, $SE = 0.002$), $\beta = -1.70$, $SE = 0.36$, $z = -4.69$, $p < .001$ (one-tailed), $OR = 0.18$, 95% CI [0.08, 0.41]. The pattern of results supports Hypotheses 1a and 1b.

For the comparison of incongruent color word condition and incongruent background color condition (Hypotheses 2a and 2b), the findings of Experiment 2 were also replicated. Responses in the incongruent color word condition were significantly slower ($M = 6.634$, $SE = 0.028$) compared to the incongruent background color condition ($M = 6.463$, $SE = 0.028$), $\beta = 0.17$, $SE = 0.02$, $t(140) = 8.18$, $p < .001$ (one-tailed), $d = 0.66$. Responses in the incongruent color word condition were also significantly less accurate ($M = .971$, $SE = 0.008$) compared

1
2
3
4
5 to the incongruent background color condition ($M = .995$, $SE = 0.002$), $\beta = -1.76$, $SE = 0.37$, z
6
7 $= -4.70$, $p < .001$ (one-tailed), $OR = 0.17$, 95% CI [0.07, 0.40]. The pattern of results supports
8
9 Hypotheses 2a and 2b.

10
11 The mismatch effect remained regardless of whether the background colors were
12 congruent with the color word or not. Responses in the incongruent color word condition
13
14 were significantly slower and less accurate than in the incongruent background color
15
16 condition. In line with Experiment 2, we found support for the assumption that color words
17
18 that are incongruent with the picture color interfere with the sentence-picture verification task
19
20 and the assumption that the perceptual simulation of color information is less influenced by
21
22 concurrently presented visual color information.
23
24
25
26
27

28 **Exploratory Analyses of Reading Time Across Experiments 1–3**

29
30 Since mental simulation is assumed to evolve during comprehension (e.g., Zwaan,
31
32 2004; Zwaan et al., 1998), we were interested in whether the mismatch effect would be
33
34 affected by the presentation of incongruent colors, depending on how deeply participants
35
36 engaged with the sentence during reading. To this end, we explored whether reading times
37
38 within each condition differentially predicted response accuracy and/or latency in the
39
40 sentence-picture verification task. In all experiments, the reading time per word was recorded
41
42 during the self-paced presentation of the experimental sentences. We computed the sum
43
44 scores per sentence (reading times per word over 4000 ms were excluded beforehand as
45
46 extreme outliers in each experiment) and included reading time as a second predictor (z -
47
48 standardized) in the GLMMs and LMMs that were used to test these hypotheses. All
49
50 conditions were included simultaneously for Experiments 2 and 3 (dummy-coded, with the
51
52 fully congruent condition as the reference category coded with 0) and main effects and the
53
54 interaction of the experimental conditions and the mean reading times were included
55
56
57
58
59
60

1
2
3
4
5 simultaneously as fixed effects. We also included random effects (random intercepts) of
6
7 participants and sentences.
8

9
10 Based on these models, we conducted simple slope analyses to estimate condition-
11 specific effects (Figure 6). The analyses revealed a significant negative slope in the
12 incongruent color word condition in Experiment 3 for response accuracy ($\beta = -0.45$, $SE =$
13 0.17 , $z = -2.63$, $p = .009$, two-tailed), suggesting that the longer participants read the
14 sentences, the worse their response accuracy. Conversely, the less time that participants
15 engaged with the sentences, the more they responded “yes” to picture colors that showed
16 visual congruence with the background color, even though the picture color was incongruent
17 to the color word, compared to participants who took more time to read the sentences. A
18 similar trend emerged in the incongruent color word condition in Experiment 2, although it
19 did not reach statistical significance ($\beta = -0.28$, $SE = 0.18$, $z = -1.51$, $p = .132$, two-tailed). In
20 both experiments, no significant effects were observed for the fully congruent or incongruent
21 background color condition (Table 1). In the incongruent condition in Experiment 1, which
22 included no experimental manipulation of background color, reading times were not
23 predictive of response accuracy ($\beta = -0.16$, $SE = 0.16$, $z = -0.99$, $p = .320$, two-tailed). This
24 finding suggests that the negative effects of reading time on accuracy in the incongruent
25 conditions of Experiment 2 and 3 were due to the processing of incongruent background
26 colors in these experiments. Analogous analyses with response latency as the dependent
27 variable showed positive slopes of reading time across all conditions (see Table A1 in the
28 Online Supplement), reflecting an overall close correlation of reading and reaction times.
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

53 **General Discussion**

54
55 In three preregistered experiments, we examined the role of perceptual simulations of
56 color in the comprehension of sentences that contained color words. Our first objective was
57
58
59
60

1
2
3
4
5 to examine whether a mismatch effect occurs when explicit color words are used in a
6
7 sentence-picture verification task. Experiment 1 yielded the expected mismatch effect,
8
9 extending previous results of sentence-picture verification tasks in the perceptual dimension
10
11 of color (de Koning et al., 2017a; Hoeben Mannaert et al., 2017, 2021; Zwaan & Pecher,
12
13 2012). The results indicate that the color of focal objects is routinely activated during
14
15 comprehension and also when color words are explicitly mentioned. Our experiments
16
17 contribute to the growing body of literature supporting the assumption that color is activated
18
19 in a similar manner to other perceptual dimensions during comprehension (see de Koning et
20
21 al., 2017a) and not in a different manner as has been discussed in the past (see Connell,
22
23 2007). Methodologically, the use of color-matching fillers ensured that the response pattern
24
25 was not based on strategic responses but instead were routine activations of sentence-related
26
27 simulations, which is in line with previous research (van Zuijlen et al., 2024). Regarding
28
29 embodied theories of language comprehension—particularly Barsalou’s (1999) theory of
30
31 perceptual symbol systems—our results support the view that perceptual representations
32
33 (specifically color representations) are routinely activated during comprehension.
34
35
36
37
38

39
40 Our second objective was to examine the functionality of perceptual simulations of
41
42 colors in sentence comprehension. Together with the experimental sentences, we presented
43
44 conflicting visual colors that we expected to impair the comprehension of color words,
45
46 resulting in altered performance in the sentence-picture verification task. This expectation
47
48 was not supported. Although color words that were incongruent with the picture color elicited
49
50 significantly slower and less accurate responses than color words that were congruent with
51
52 the picture color (corresponding to typical mismatch effect), this effect was not affected by
53
54 the parallel processing of (conflicting) visual background color information. The results were
55
56 supported in Experiment 3, a direct replication study with an independent sample.
57
58
59
60

1
2
3
4
5 We derive two main conclusions from the results. First, the mismatch effect found in
6
7 Experiment 1 appeared to be robust across all experiments and was not affected by whether
8
9 the visually presented colors were congruent with the color words or not. This finding
10
11 indicates that color representations of focal objects are routinely activated during
12
13 comprehension when their colors are explicitly mentioned, extending previous research on
14
15 the continuous activation of color representations (Hoeben Mannaert et al., 2021). Second,
16
17 the results suggest that the comprehension of color information is less influenced by
18
19 concurrently presented visual color information than would have been expected based on the
20
21 results of neuroimaging studies that have identified common processing areas for simulated
22
23 and perceived color (Simmons et al., 2007; Van Dam et al., 2012) and studies that have
24
25 demonstrated that the retrieval of visual information can be influenced by visual stimuli
26
27 (Davis et al., 2020; Edmiston & Lupyan, 2017; Ostarek & Huettig, 2017). Based on these
28
29 studies, we had assumed that the simulation of verbal color information can be influenced by
30
31 visual color information when simulation is functionally involved in comprehension
32
33 processes. However, this assumption was not supported by our results. Instead, our results are
34
35 in line with findings by Ostarek et al. (2019) who manipulated the sentence-picture
36
37 verification task by presenting visual shapes (“visual noise”) and irrelevant objects
38
39 (“semantic noise”) during auditory sentence presentation. Visual noise did not affect the
40
41 mismatch effect, whereas semantic noise did. Ostarek et al. concluded that visual processes
42
43 are not necessarily involved in the cognitive mechanisms underlying the mismatch effect and
44
45 that this effect may therefore not be (primarily) grounded in modal simulations. In a similar
46
47 approach, Buchner et al. (2025) examined the influence of direction-related perceptual
48
49 stimuli on the processing of direction-related narrative stories and likewise found no match or
50
51 mismatch effects. Consistent with these findings, we also found no evidence for a modulatory
52
53 effect of visual processes – specifically, the processing of concurrently presented colors in the
54
55
56
57
58
59
60

1
2
3
4
5 present study – on the mismatch effect for colors. It may therefore be speculated that
6
7 perceptual simulations are rather a “by-product” of comprehension (Strozyk et al., 2019, p.
8
9 407; for a theoretical discussion, see Goldinger et al., 2016).

10
11 An alternative interpretation of the unexpected results is offered by theories of
12
13 conflict processing. According to the conflict monitoring theory proposed by Botvinick et al.
14
15 (2001), information processing is continuously monitored to detect potential conflicts
16
17 between competing response options (e.g., as in Stroop tasks). When such a conflict is
18
19 detected, cognitive control is upregulated, leading to heightened attention and more accurate
20
21 responses. In line with this framework, and contrary to our hypotheses in Experiment 2, that
22
23 parallel processing of conflicting background colors would reduce response accuracy and
24
25 increase response latency, the simultaneous visual presentation of conflicting colors may
26
27 instead have enhanced attention and comprehension, thereby accounting for the absence of an
28
29 interference effect in Experiment 2.
30
31
32
33

34 Although no measurable direct effects of visual background color on response latency
35
36 and accuracy were found, exploratory analyses revealed differences between the conditions
37
38 when considering reading time as an indirect index of processing effort during sentence
39
40 reading. In the incongruent color word condition, when color word and picture color were
41
42 incongruent, but picture color was concurrently congruent with the background color, the
43
44 mismatch interference was reduced (i.e., response accuracy was less deteriorated) for
45
46 participants who read faster compared to participants who read more slowly. Assuming that
47
48 shorter reading times indicate more superficial reading processes and longer reading times
49
50 indicate deeper reading processes, participants who processed the sentences more deeply
51
52 might have been more likely to be affected by the incongruency between color word and
53
54 picture color and less likely to be affected by the background color because they were
55
56 engaged in a thorough simulation of the sentence content, resulting in a comparatively
57
58
59
60

1
2
3
4
5 stronger mismatch interference. The exploratory findings therefore suggest that superficial
6
7 reading processes are more strongly influenced by concurrently presented colors than deep
8
9 reading processes, whereas deep reading processes might be accompanied by more resistant
10
11 simulations. In addition, the conflict monitoring theory proposed by Botvinick et al. (2001)
12
13 can again be applied to explain that participants who read more slowly may have perceived
14
15 the color conflict, thereby engaging higher cognitive control, in contrast to those who spent
16
17 less time on the sentence.
18
19

20
21 As the incongruent color word condition is the only one involving a word–picture
22
23 mismatch, an important alternative interpretation is that the reading time effect might reflect
24
25 the perceived semantic incongruence instead of the interference of background color. It
26
27 seems reasonable that longer reading times (or deeper processing) are associated with
28
29 increasing salience of semantic incongruence. At this point, the lack of a fully crossed design
30
31 prevents a clear interpretation of this effect.
32
33

34
35 However, the effect did not occur in Experiment 1, which also required responses to
36
37 incongruent picture colors but with no experimental manipulation in the form of background
38
39 colors, making this explanation unlikely. In sum, the exploratory findings on reading time
40
41 support the interpretation that incongruent background colors can affect perceptual
42
43 simulations in sentence comprehension but only when readers process the sentence
44
45 superficially. Given that this interpretation is based on exploratory findings, it remains to be
46
47 tested in confirmatory studies.
48
49

50 51 **Limitations and Future Directions**

52
53 One limitation of our findings is that our experimental designs were reduced to keep
54
55 the procedure short and avoid habituation or practice effects. The downside of the reduced
56
57 design is that our method lacked a fully crossed design of congruent/incongruent background
58
59 color and congruent/incongruent object color. Such a design might have enabled more
60

1
2
3
4
5 comprehensive conclusions. A second limitation is that despite this precaution, we cannot
6
7 rule out the possibility that participants became habituated to the repeated presentation of
8
9 visual colors. Even in simple Stroop tasks, habituation effects have been observed following
10
11 frequent task repetition (see Stroop, 1935, Experiment 3). Therefore, habituation effects
12
13 might have prevented the effects of background color from occurring. A third limitation is
14
15 that the reading-time dependent effect of congruent vs. incongruent background color on the
16
17 match effect was observed only for response accuracy, not response latency, and was based
18
19 on exploratory analyses, requiring confirmatory testing in future studies.
20
21
22

23
24 Future research should therefore further investigate the role of visual interference
25
26 during on-line reading, depending on the depth of sentence processing, as this focus remains
27
28 a promising yet underexplored avenue for understanding the mental processes involved in
29
30 reading. Further studies are required to investigate the role of simulations in various
31
32 perceptual areas and to draw more far-reaching conclusions about their significance in
33
34 language comprehension.
35
36

37 **Conclusion**

38
39 Our results indicate that the mismatch effect for sentences containing color words
40
41 remains robust even when an incongruent background color is presented concurrently during
42
43 reading. This finding suggests that the comprehension of explicit verbal color information is
44
45 less susceptible to interference from simultaneously presented visual color information than
46
47 might be expected based on findings from neuroimaging studies. Overall, the results of the
48
49 present experiments provide evidence that color representations are routinely activated during
50
51 language comprehension when colors are explicitly mentioned.
52
53
54
55
56
57
58
59
60

Supplementary Material

The Supplementary Material is available at: qjep.sagepub.com

Data Accessibility Statement

The data and materials from the present experiment are publicly available at the Open Science Framework website:

https://osf.io/7h2py/?view_only=f64033c9cd6a4d8abd2c8c33e6124ca2

Acknowledgements

We would like to thank our assistants who helped with data collection and Juliana Gollwitzer for her help in creating the stimuli.

Statements and Declarations

Ethical considerations

We have not applied for ethical approval for the experiments, as according to the ethical guidelines of the German Society for Psychology (DGPs) and the local ethics committee, prior review by an Institutional Review Board is not mandatory for research that provides signed informed consent from participants, collects data anonymously, and has no foreseeable negative impact on participants.

Consent to participate

Informed consent was obtained from all participants before the start of the experiment. Information was presented on the screen and consent was recorded via Inquisit.

Consent for publication

No case reports, images or similar included, therefore not applicable.

Declaration of conflicting interest

1
2
3
4
5 The authors declared no potential conflicts of interest with respect to the research,
6
7 authorship and/or publication of this article.
8

9 *Funding statement*

10
11 We received funding for this research from the Human Dynamics Center of the
12
13 Faculty of Human Sciences at the University of Würzburg.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Peer Review Version

References

- 1
2
3
4
5
6
7 Andrade, J., May, J., Deeprase, C., Baugh, S.-J., & Ganis, G. (2014). Assessing vividness of
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Andrade, J., May, J., Deeprase, C., Baugh, S.-J., & Ganis, G. (2014). Assessing vividness of
mental imagery: The Plymouth Sensory Imagery Questionnaire. *British Journal of
Psychology*, *105*(4), 547–563. <https://doi.org/10.1111/bjop.12050>
- Azañón, E., Pounder, Z., Figueroa, A., & Reeder, R. R. (2025). Individual variability in
mental imagery vividness does not predict perceptual interference with imagery: A
replication study of Cui et al. (2007). *Journal of Experimental Psychology: General*,
154(7), 2043–2057. <https://doi.org/10.1037/xge0001756>
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4),
577–660. <https://doi.org/10.1017/S0140525X99002149>
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645.
<https://doi.org/10.1146/annurev.psych.59.103006.093639>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed
random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–
412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). Fitting linear mixed-effects models
using lme4. *Journal of Statistical Software*, *67*(1), 1–48.
<https://doi.org/10.18637/jss.v067.i01>.
- Berndt, E., Dudschig, C., & Kaup, B. (2020). Green as a cbemcuru: Modal as well as amodal
color cues can help to solve anagrams. *Psychological Research*, *84*(2), 491–501.
<https://doi.org/10.1007/s00426-018-1055-y>
- Birch, J. (1997). Efficiency of the Ishihara test for identifying red-green colour deficiency.
Ophthalmic and Physiological Optics, *17*(5), 403–408. [https://doi.org/10.1016/S0275-5408\(97\)00022-7](https://doi.org/10.1016/S0275-5408(97)00022-7)

- 1
2
3
4
5 Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict
6
7 monitoring and cognitive control. *Psychological Review*, *108*(3), 624–
8
9 652. <https://doi.org/10.1037/0033-295X.108.3.624>
- 11 Buchner, E. L., Lenhard, W., & Richter, T. (2025). Comprehending dynamic directions: the
12
13 role of perceptual simulations. *Journal of Cognitive Psychology*, *37*(8), 807–823.
14
15
16 <https://doi.org/10.1080/20445911.2025.2546865>
- 18 Connell, L. (2007). Representing object colour in language comprehension. *Cognition*,
19
20
21 *102*(3), 476–485. <https://doi.org/10.1016/j.cognition.2006.02.009>
- 23 Connell, L., & Lynott, D. (2009). Is a bear white in the woods? Parallel representation of
24
25 implied object color during language comprehension. *Psychonomic Bulletin &*
26
27 *Review*, *16*, 573–577. <https://doi.org/10.3758/PBR.16.3.573>
- 30 Cui, X., Jeter, C. B., Yang, D., Montague, P. R., & Eagleman, D. M. (2007). Vividness of
31
32 mental imagery: Individual variability can be measured objectively. *Vision*
33
34
35 *Research*, *47*(4), 474–478. <https://doi.org/10.1016/j.visres.2006.11.013>
- 37 Davis, C. P., Joergensen, G. H., Boddy, P., Dowling, C., & Yee, E. (2020). Making it harder
38
39 to “see” meaning: The more you see something, the more its conceptual
40
41 representation is susceptible to visual interference. *Psychological Science*, *31*(5), 505–
42
43 517. <https://doi.org/10.1177/0956797620910748>
- 46 de Koning, B. B., Wassenburg, S. I., Bos, L. T., & van der Schoot, M. (2017a). Mental
47
48 simulation of four visual object properties: Similarities and differences as assessed by
49
50 the sentence–picture verification task. *Journal of Cognitive Psychology*, *29*(4), 420–
51
52 432. <https://doi.org/10.1080/20445911.2017.1281283>
- 55 de Koning, B. B., Wassenburg, S. I., Bos, L. T., & van der Schoot, M. (2017b). Size does
56
57 matter: Implied object size is mentally simulated during language
58
59
60

- 1
2
3
4
5 comprehension. *Discourse Processes*, 54(7), 493–503.
6
7 <http://dx.doi.org/10.1080/0163853X.2015.1119604>
8
9 del Prado Martín, F. M., Hauk, O., & Pulvermüller, F. (2006). Category specificity in the
10
11 processing of color-related and form-related words: An ERP
12
13 study. *Neuroimage*, 29(1), 29–37. <https://doi.org/10.1016/j.neuroimage.2005.07.055>
14
15
16 Edmiston, P., & Lupyan, G. (2017). Visual interference disrupts visual knowledge. *Journal of*
17
18 *Memory and Language*, 92, 281–292. <https://doi.org/10.1016/j.jml.2016.07.002>
19
20
21 Engelen, J. A., Bouwmeester, S., de Bruin, A. B., & Zwaan, R. A. (2011). Perceptual
22
23 simulation in developing language comprehension. *Journal of Experimental Child*
24
25 *Psychology*, 110(4), 659–675. <https://doi.org/10.1016/j.jecp.2011.06.009>
26
27
28 Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical
29
30 power analysis program for the social, behavioral, and biomedical sciences. *Behavior*
31
32 *Research Methods*, 39, 175–191. <https://doi.org/10.3758/bf03193146>
33
34
35 Glaser, W. R. (1992). Picture naming. *Cognition*, 42(1–3), 61–105.
36
37 [https://doi.org/10.1016/0010-0277\(92\)90040-O](https://doi.org/10.1016/0010-0277(92)90040-O)
38
39
40 Goldinger, S. D., Papesh, M. H., Barnhart, A. S., Hansen, W. A., & Hout, M. C. (2016). The
41
42 poverty of embodied cognition. *Psychonomic Bulletin & Review*, 23, 959–978.
43
44 <https://doi.org/10.3758/s13423-015-0860-1>
45
46
47 Hoeben Mannaert, L. N., Dijkstra, K., & Zwaan, R. A. (2017). Is color an integral part of a
48
49 rich mental simulation? *Memory & Cognition*, 45, 974–982.
50
51 <https://doi.org/10.3758/s13421-017-0708-1>
52
53
54 Hoeben Mannaert, L. N., Dijkstra, K., & Zwaan, R. A. (2021). Is color continuously activated
55
56 in mental simulations across a broader discourse context? *Memory & Cognition*, 49,
57
58 127–147. <https://doi.org/10.3758/s13421-020-01078-6>
59
60

- 1
2
3
4
5 Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not)
6
7 and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.
8
9 <https://doi.org/10.1016/j.jml.2007.11.007>
10
11 Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A.
12
13 A., & Zwaan, R. A. (2005). Perception of motion affects language processing.
14
15 *Cognition*, *94*(3), 79–89. <https://doi.org/10.1016/j.cognition.2004.06.005>
16
17 Kaschak, M. P., Zwaan, R. A., Aveyard, M., & Yaxley, R. H. (2006). Perception of auditory
18
19 motion affects language processing. *Cognitive Science*, *30*(4), 733–744.
20
21 https://doi.org/10.1207/s15516709cog0000_54
22
23
24
25 Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery.
26
27 *Nature Reviews Neuroscience*, *2*(9), 635–642. <https://doi.org/10.1038/35090055>
28
29
30 Koster, D., Cadierno, T., & Chiarandini, M. (2019). Mental simulation of object orientation
31
32 and size: A conceptual replication with second language learners. *Journal of the*
33
34 *European Second Language Association*, *2*(1), 38–48.
35
36 <https://doi.org/10.22599/jesla.39>
37
38
39 Lenth, R. (2025). emmeans: Estimated marginal means, aka least-squares means.
40
41 <https://doi.org/10.32614/CRAN.package.emmeans>
42
43
44 Liu, D., Wang, L., & Han, Y. (2024). Mental simulation of colour properties during language
45
46 comprehension: Influence of context and comprehension stages. *Cognitive*
47
48 *Processing*, *25*(4), 587–600. <https://doi.org/10.1007/s10339-024-01201-4>
49
50
51 MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative
52
53 review. *Psychological Bulletin*, *109*(2), 163–203.
54
55
56 Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis
57
58 and a new proposal for grounding conceptual content. *Journal of Physiology – Paris*,
59
60 *102*(1–3), 59–70. <https://doi.org/10.1016/j.jphysparis.2008.03.004>

- 1
2
3
4
5 Mak, M., & Willems, R. M. (2021). Mental simulation during literary reading. In D. Kuiken
6
7 & A. M. Jacobs (Eds.), *Handbook of empirical literary studies* (pp. 63–84). De
8
9 Gruyter. <https://doi.org/10.1515/9783110645958-004>
10
11
12 Millisecond. (2021). *Inquisit (Version 6.6.1)*. [Computer software]. Millisecond Software.
13
14 Ostarek, M., & Bottini, R. (2021). Towards strong inference in research on embodiment–
15
16 possibilities and limitations of causal paradigms. *Journal of Cognition*, 4(1), 1–21.
17
18 <https://doi.org/10.5334/joc.139>
19
20
21 Ostarek, M., & Huettig, F. (2017). A task-dependent causal role for low-level visual
22
23 processes in spoken word comprehension. *Journal of Experimental Psychology:*
24
25 *Learning, Memory, and Cognition*, 43(8), 1215–1224.
26
27 <https://doi.org/10.1037/xlm0000375>
28
29
30 Ostarek, M., & Huettig, F. (2019). Six challenges for embodiment research. *Current*
31
32 *Directions in Psychological Science*, 28(6), 593–599.
33
34 <https://doi.org/10.1177/0963721419866441>
35
36
37 Ostarek, M., Joosen, D., Ishag, A., De Nijs, M., & Huettig, F. (2019). Are visual processes
38
39 causally involved in “perceptual simulation” effects in the sentence-picture
40
41 verification task? *Cognition*, 182, 84–94.
42
43 <https://doi.org/10.1016/j.cognition.2018.08.017>
44
45
46 Pulvermüller, F. (1999). Words in the brain’s language. *Behavioral and Brain Sciences*,
47
48 22(2), 253–279. <https://doi.org/10.1017/S0140525X9900182X>
49
50
51 Pulvermüller, F., & Hauk, O. (2006). Category-specific conceptual processing of color and
52
53 form in left fronto-temporal cortex. *Cerebral Cortex*, 16(8), 1193–1201.
54
55 <https://doi.org/10.1093/cercor/bhj060>
56
57
58 Richter, T., & van Holt, N. (2005). ELVES: Ein computergestütztes Diagnostikum zur
59
60 Erfassung der Effizienz von Teilprozessen des Leseverstehens [ELVES: A computer-

- 1
2
3
4
5 aided diagnostic tool for recording the efficiency of the sub-processes of reading
6
7 comprehension]. *Diagnostica*, 51, 169–182. <https://doi.org/10.1026/0012->
8
9 1924.51.4.169
10
11 Richter, T., & Zwaan, R. A. (2009). Processing of color words activates color
12
13 representations. *Cognition*, 111, 383–389.
14
15 <https://doi.org/10.1016/j.cognition.2009.02.011>
16
17
18 Simmons, W. K., Ramjee, V., Beauchamp, M. S., McRae, K., Martin, A., & Barsalou, L. W.
19
20 (2007). A common neural substrate for perceiving and knowing about color.
21
22 *Neuropsychologia*, 45(12), 2802–2810.
23
24 <https://doi.org/10.1016/j.neuropsychologia.2007.05.002>
25
26
27 Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from
28
29 verbal context on picture recognition. *Psychological Science*, 12(2), 153–156.
30
31 <https://doi.org/10.1111/1467-9280.00326>
32
33
34 Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of*
35
36 *Experimental Psychology*, 18(6), 643–662. <https://doi.org/10.1037/h0054651>
37
38
39 Strozyk, J. V., Dudschig, C., & Kaup, B. (2019). Do I need to have my hands free to
40
41 understand hand-related language? Investigating the functional relevance of
42
43 experiential simulations. *Psychological Research*, 83, 406–418.
44
45 <https://doi.org/10.1007/s00426-017-0900-8>
46
47
48 Suggate, S. P. (2024). Beyond self-report: Measuring visual, auditory, and tactile mental
49
50 imagery using a mental comparison task. *Behavior Research Methods*, 56(8), 8658–
51
52 8676. <https://doi.org/10.3758/s13428-024-02496-z>
53
54
55 Van Dam, W. O., Van Dijk, M., Bekkering, H., & Rueschemeyer, S. A. (2012). Flexibility in
56
57 embodied lexical-semantic representations. *Human Brain Mapping*, 33(10), 2322–
58
59 2333. <https://doi.org/10.1002/hbm.21365>
60

- 1
2
3
4
5 van Zuijlen, S. J., Singh, S., Gunawan, K., Pecher, D., & Zeelenberg, R. (2024). Automatic
6
7 mental simulation in native and non-native speakers. *Memory & Cognition*, *52*(5),
8
9 1152–1163. <https://doi.org/10.3758/s13421-024-01533-8>
10
11
12 Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language
13
14 comprehension. In B. H. Ross (Ed.), *The psychology of learning and motivation*. (Vol.
15
16 44). Academic Press: New York.
17
18
19 Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words:
20
21 Dynamic representations in language comprehension. *Cognitive Science*, *28*(4), 611–
22
23 619. https://doi.org/10.1207/s15516709cog2804_5
24
25
26 Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension:
27
28 Six replication attempts. *PloS One*, *7*(12), e51382.
29
30 <https://doi.org/10.1371/journal.pone.0051382>
31
32
33 Zwaan, R. A., Radvansky, G. A., Hilliard, A. E., & Curiel, J. M. (1998). Constructing
34
35 multidimensional situation models during reading. *Scientific Studies of Reading*, *2*(3),
36
37 199–220. https://doi.org/10.1207/s1532799xssr0203_2
38
39
40 Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally
41
42 represent the shapes of objects. *Psychological Science*, *13*(2), 168–171.
43
44 <https://doi.org/10.1111/1467-9280.00430>
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure Captions**Figure 1**

Example of the Trial Sequence in Experiment 1

[Figure 1.jpg]

Note. The same experimental sentence was combined with a congruent picture (“Congruent”) and an incongruent picture (“Incongruent”) but only appeared once per list and per participant. The sentences were presented in German. The English translations of the sentences may not be exact, as they were intended to contain seven words like the German originals (original versions: “Der große grüne LKW steht am Straßenrand” and “Der neue grüne Besen steht griffbereit da”). The correct answer for the experimental trials (congruent/incongruent) was always “Yes”. The correct answer for filler trials was always “No”. RA = response accuracy. RL = response latency.

Figure 2

Mean Response Latency and Response Accuracy per Condition in Experiment 1

[Figure 2.jpg]

Note. Mean response latency (only correct responses) (A) and mean response accuracy (B) for congruent and incongruent sentence-picture verification tasks. Error bars represent standard errors of the mean. ** $p < .01$, *** $p < .001$.

Figure 3

Example of the Trial Sequence in Experiment 2 and 3

[Figure 3.jpg]

1
2
3
4
5 *Note.* Each experimental sentence existed in three versions but only appeared once per list
6
7 and per participant. The sentences were presented in German. The English translations of the
8
9 sentences may not be exact, as they were intended to contain seven words like the German
10
11 originals (original versions: “Der große grüne LKW steht am Straßenrand” and “Der neue
12
13 grüne Besen steht griffbereit da”). The correct answer for all experimental trials (fully
14
15 congruent/incongruent background color/incongruent color word) was always “Yes”. The
16
17 correct answer for filler trials was always “No”. RA = response accuracy. RL = response
18
19 latency.
20
21
22
23
24

25 **Figure 4**

26
27 *Mean Response Latency and Response Accuracy per Condition in Experiment 2*

28
29 [Figure 4.jpg]

30
31
32 *Note.* Mean response latency (only correct responses) (A) and mean response accuracy (B)
33
34 for fully congruent, incongruent background color and incongruent color word sentence-
35
36 picture verification tasks. Error bars represent standard errors of the mean. *** $p < .001$.
37
38
39
40

41 **Figure 5**

42
43 *Mean Response Latency and Response Accuracy per Condition in Experiment 3*

44
45 [Figure 5.jpg]

46
47
48 *Note.* Mean response latency (only correct responses) (A) and mean response accuracy (B)
49
50 for the fully congruent, incongruent background color and incongruent color word sentence-
51
52 picture verification tasks. Error bars represent standard errors of the mean. *** $p < .001$.
53
54
55
56

57 **Figure 6**

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

*Estimated Slopes of Reading Time Predicting the Probability of a Correct Response for
Experiment 1 (A), Experiment 2 (B), and Experiment 3 (C) per Condition*

[Figure 6.jpg]

Note. Gray areas represent 95% CI per condition

Peer Review Version



Figure 1. Example of the Trial Sequence in Experiment 1

Note. The same experimental sentence was combined with a congruent picture ("Congruent") and an incongruent picture ("Incongruent") but only appeared once per list and per participant. The sentences were presented in German. The English translations of the sentences may not be exact, as they were intended to contain seven words like the German originals (original versions: "Der große grüne LKW steht am Straßenrand" and "Der neue grüne Besen steht griffbereit da"). The correct answer for the experimental trials (congruent/incongruent) was always "Yes". The correct answer for filler trials was always "No". RA = response accuracy. RL = response latency.

338x190mm (96 x 96 DPI)

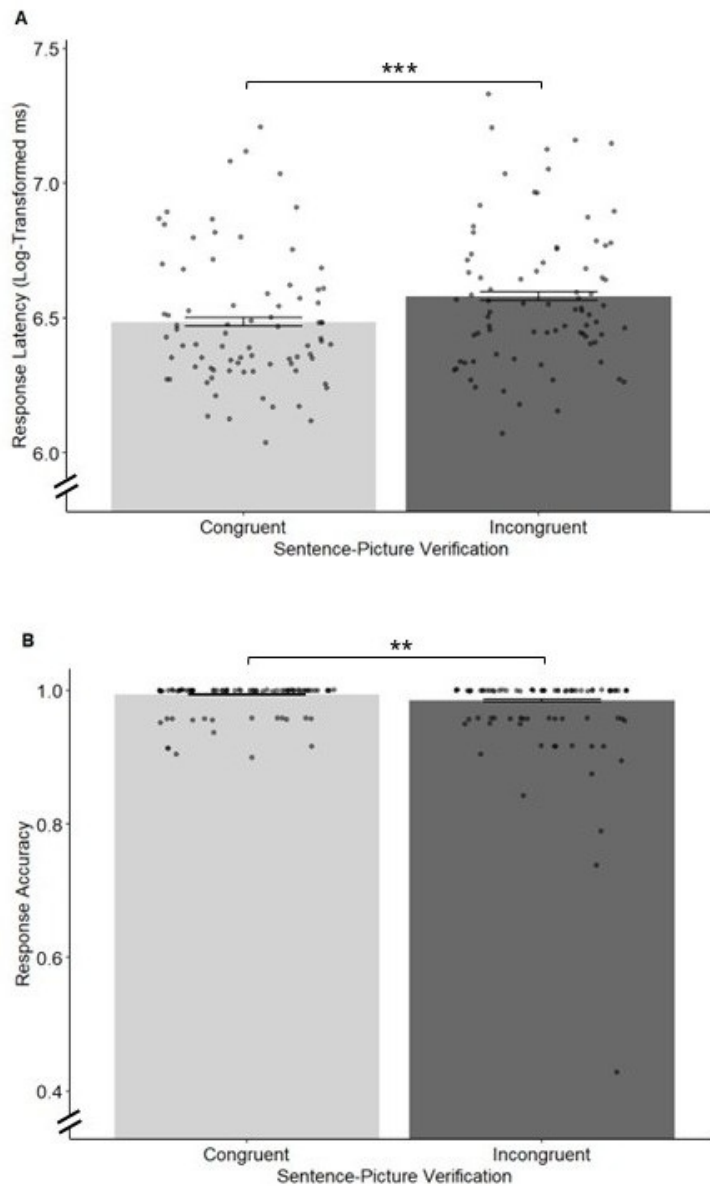


Figure 2. Mean Response Latency and Response Accuracy per Condition in Experiment 1
Note. Mean response latency (only correct responses) (A) and mean response accuracy (B) for congruent and incongruent sentence-picture verification tasks. Error bars represent standard errors of the mean. ** $p < .01$, *** $p < .001$.

127x190mm (96 x 96 DPI)

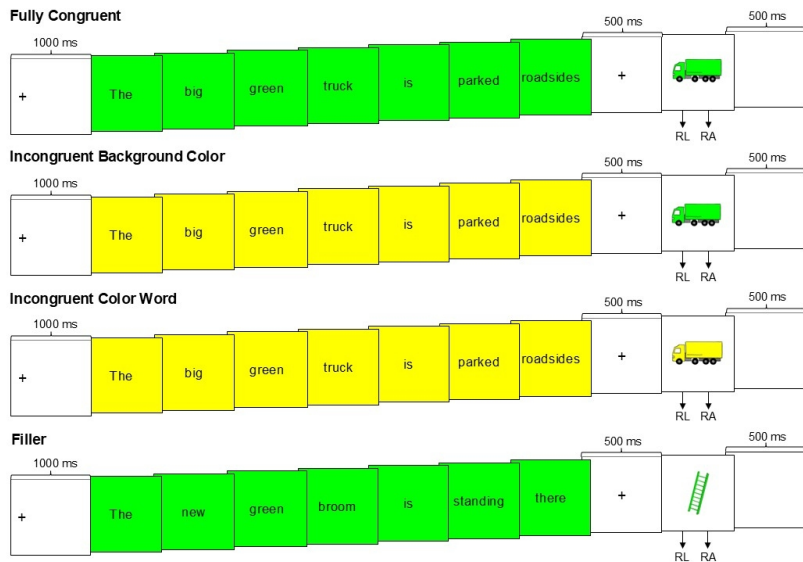


Figure 3. Example of the Trial Sequence in Experiment 2 and 3

Note. Each experimental sentence existed in three versions but only appeared once per list and per participant. The sentences were presented in German. The English translations of the sentences may not be exact, as they were intended to contain seven words like the German originals (original versions: "Der große grüne LKW steht am Straßenrand" and "Der neue grüne Besen steht griffbereit da"). The correct answer for all experimental trials (fully congruent/incongruent background color/incongruent color word) was always "Yes". The correct answer for filler trials was always "No". RA = response accuracy. RL = response latency.

338x190mm (96 x 96 DPI)

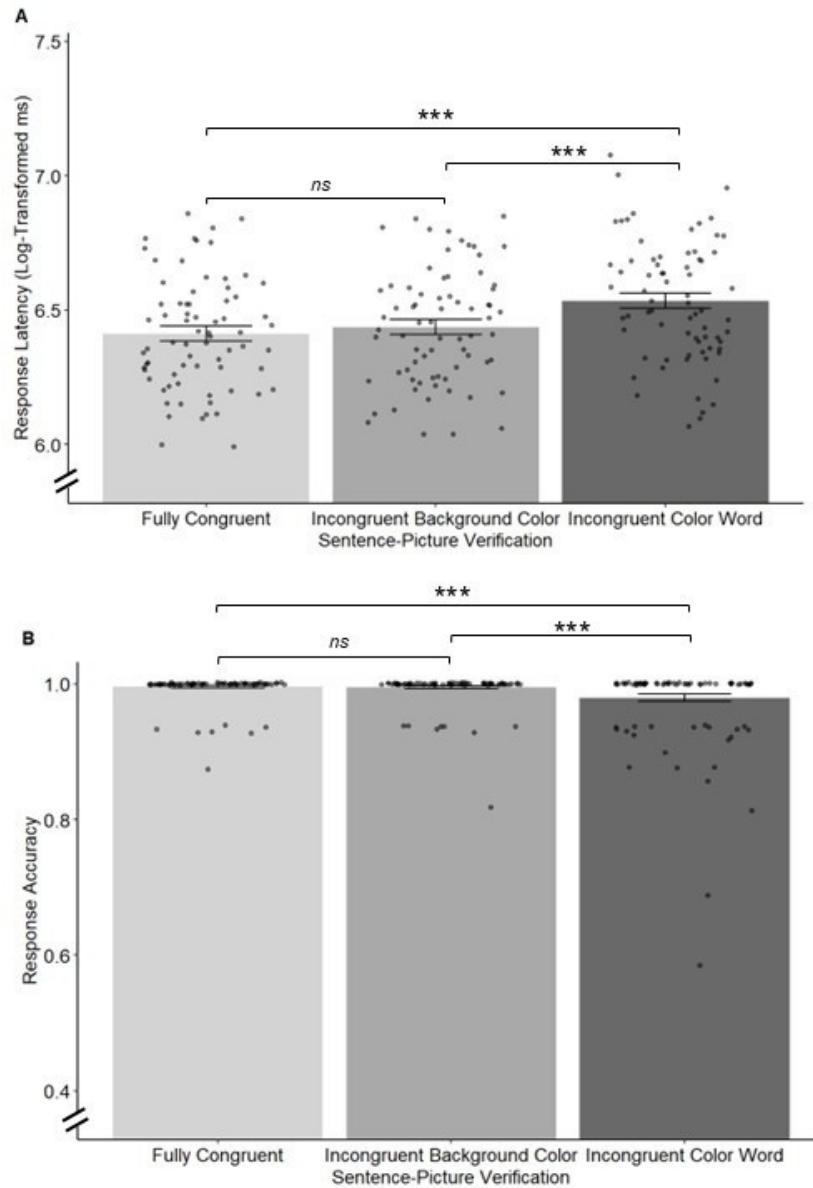


Figure 4. Mean Response Latency and Response Accuracy per Condition in Experiment 2
Note. Mean response latency (only correct responses) (A) and mean response accuracy (B) for fully congruent, incongruent background color and incongruent color word sentence-picture verification tasks. Error bars represent standard errors of the mean. *** $p < .001$.

133x190mm (96 x 96 DPI)

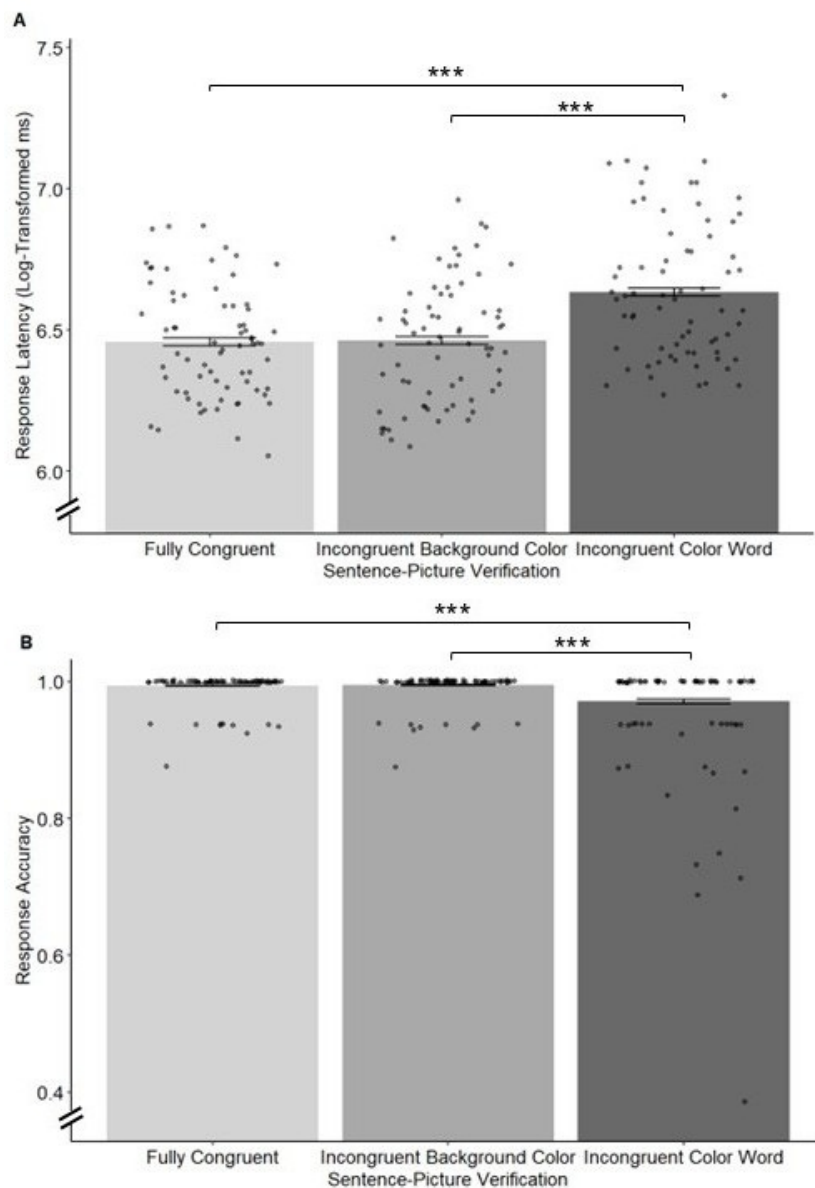


Figure 5. Mean Response Latency and Response Accuracy per Condition in Experiment 3
 Note. Mean response latency (only correct responses) (A) and mean response accuracy (B) for the fully congruent, incongruent background color and incongruent color word sentence-picture verification tasks. Error bars represent standard errors of the mean. *** $p < .001$.

139x190mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

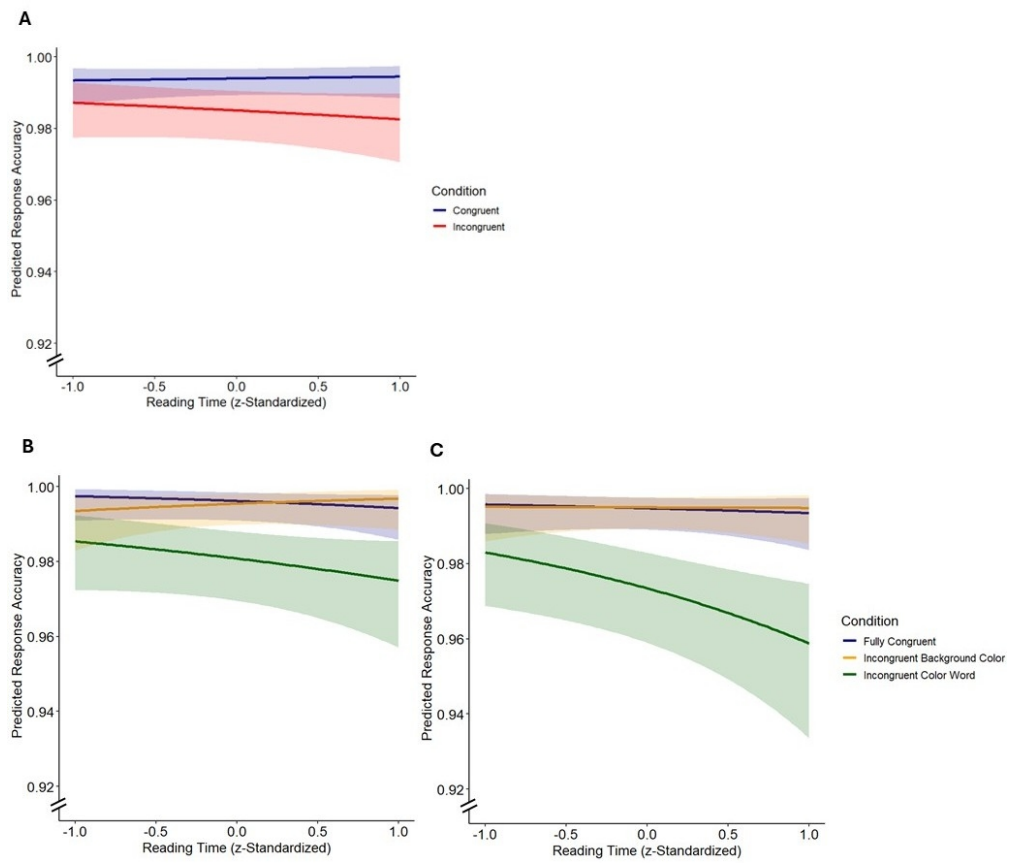


Figure 6. Estimated Slopes of Reading Time Predicting the Probability of a Correct Response for Experiment 1 (A), Experiment 2 (B), and Experiment 3 (C) per Condition
Note. Gray areas represent 95% CI per condition

273x237mm (96 x 96 DPI)

Processing of Color Words

Table 1

Simple Slopes of Reading Time on Response Accuracy per Condition Across All Experiments

Condition per Experiment	β	<i>SE</i>	<i>z</i>	<i>p</i>
Experiment 1				
Congruent	0.09	0.20	0.44	.663
Incongruent	-0.16	0.16	-0.99	.320
Experiment 2				
Fully Congruent	-0.41	0.33	-1.23	.219
Incongruent Background Color	0.36	0.38	0.94	.350
Incongruent Color Word	-0.28	0.18	-1.51	.132
Experiment 3				
Fully Congruent	-0.22	0.31	-0.70	.483
Incongruent Background Color	-0.03	0.34	-0.09	.932
Incongruent Color Word	-0.45	0.17	-2.63	.009

Note. All *p* values are based on two-sided tests. Reading time was z-standardized.