

PSYCHONOMIC BULLETIN & REVIEW

Enhancing visual search efficiency through cross-modal priming with auditory primes

Journal:	<i>Psychonomic Bulletin & Review</i>
Manuscript ID	Draft
Manuscript Type:	Brief Report
Date Submitted by the Author:	n/a
Complete List of Authors:	Wentura, Dirk; Universität des Saarlandes, Psychology Rohr, Michaela; Universität des Saarlandes, Psychology Folyi, Tímea; Universität des Saarlandes, Psychology
Keywords:	priming, visual search, cross-modal attention, working memory

SCHOLARONE™
Manuscripts



UNIVERSITÄT
DES
SAARLANDES

Universität des Saarlandes · Postfach · D-66041 Saarbrücken

James R. Brockmole
Editor Psychonomic Bulletin & Review

via PBR Submission Portal

Fakultät für Empirische Humanwissenschaften und
Wirtschaftswissenschaft, Fachrichtung Psychologie

Allgemeine Psychologie und
Methodenlehre

Univ.-Prof.
Dr. Dirk Wentura

Gebäude A2 4
Camous A2 4
D-66041 Saarbrücken

Telefon: 0 681 · 302 4642
Telefax: 0 681 · 302 4049

E-Mail: wentura@
mx.uni-saarland.de

Saarbrücken, Jan 24th 2020

■ **Submission to Psychonomic Bulletin & Review**

Dear Dr. Brockmole,

attached you will find a manuscript entitled "Enhancing visual search efficiency through cross-modal priming with auditory primes" that I, Michaela Rohr, and Tímea Folyi are submitting to *Psychonomic Bulletin & Review*.

The material has not been published and is not under consideration for publication elsewhere.

Thank you for considering our article.

Sincerely,

Dirk Wentura

RUNNING HEAD: CROSS-MODAL ENHANCEMENT OF SEARCH EFFICACY

Enhancing visual search efficiency through cross-modal priming with auditory
primes

Dirk Wentura • Michaela Rohr

Timea Folyi

Saarland University, Saarbrücken, Germany

Word count: 2464 (sum of Introduction, Results, and Discussion sections)

Dirk Wentura
Saarland University
Faculty of Behavioral Sciences
Department of Psychology
Building A2 4
D-66041 Saarbrücken
e-mail: wentura@mx.uni-saarland.de.

Abstract

In the present experiment, we tested whether auditory primes can enhance attention to and thereby processing of subsequently appearing visual targets in a visual search task. Indeed, we found that task-irrelevant, time-compressed spoken auditory color words led to faster responses to color targets in a visual search task, replicating existing priming effects (Mahr & Wentura, 2014). Crucially, by varying set size, we provide direct evidence for an enhancement of search efficiency by these primes, that is, a flatter search slope for congruent relative to incongruent trials. The results are interpreted in terms of recent theorizing in working memory research: The primed item has privileged status as the focus of attention and therefore the potential to guide visual attention.

Keywords: priming, visual search, working memory, cross-modal

Enhancing visual search efficiency through cross-modal priming with auditory primes

Does a brief and spatially uninformative auditory prime facilitate visual search for a semantically related item? For example, does the spoken word “red” facilitate search for a red target color-patch in a display that contains several other distracting color patches? Or, to take an everyday example: Will a cat be more easily detected in a complex visual environment if a (spatially uninformative) “meow” accompanies the visual scene?

There are several studies by Iordanescu et al. (Iordanescu, Grabowecky, Franconeri, Theeuwes, & Suzuki, 2010; Iordanescu, Grabowecky, & Suzuki, 2011; Iordanescu, Guzman-Martinez, Grabowecky, & Suzuki, 2008; see also Zweig, Suzuki, & Grabowecky, 2015) that addressed crossmodal visual search. For example, Iordanescu et al. (2008) presented in each trial four pictures of objects on the screen. The designated target (e.g., cat) had to be localized. Synchronously to the search display a sound was presented that was either characteristic for the target object (e.g., a “meow”) or not. Indeed, a target-related sound speeded up responses. Iordanescu et al. (2010) conceptually replicated this effect with saccadic search times, which were lower if the target-related sound accompanied the visual display.

However, in basic visual search experiments, the marker of search efficiency is the steepness of the search slope, that is, the increase in reaction time per additional distractor item (see, e.g., Wolfe, 1998). A search is considered more efficient if slope values are smaller. To calculate slopes, however, it is indispensable to vary set size (i.e., the number of items presented in a trial) across trials. This was not realized in the experiments by Iordanescu et al. (2008) and Iordanescu et al. (2010). However, in an

1
2
3
4 additional experiment, Iordanescu et al. (2008) used an eight-items display (instead of
5 four items as in the experiment reported above). If one calculates search slopes across-
6
7 experiments, there is no evidence for increased efficiency by target-related auditory
8
9 primes. Knoeferle, Knoeferle, Velasco, and Spence (2016) used the paradigm developed
10
11 by Iordanescu et al. (2008) in an applied setting. In Experiment 4B, participants' task
12
13 was to search for a brand in one versus three rows of six products each, similar to a set
14
15 size variation manipulation. They found significant facilitation of a brand primed by a
16
17 related auditory prime only in the three rows condition, a result that can be seen (with
18
19 some caution) as equivalent to increased search efficiency in the related priming
20
21 condition. However, due to the applied context, the experiment was a single trial online
22
23 study (i.e., the set size variation was between participants) with an *amazon*-style product
24
25 selection page as display, making it difficult to relate the result to typical visual search
26
27 experiments.¹
28
29
30
31
32
33

34 Mahr and Wentura (2014) approached the test of the basic hypothesis from a
35
36 somewhat different route. They conducted cross-modal Stroop experiments with visual
37
38 targets (i.e., colored circles) and auditory primes (i.e., spoken color words). Overall a
39
40 large congruency effect was found, that is, responses were considerably faster if the
41
42 auditory prime matched the visual target color compared to incongruent pairings (see
43
44 Mahr & Wentura, 2018, for a conceptual replication with simple visual icons and
45
46 corresponding verbal labels). Notably, primes were either uncompressed (i.e., 400 ms
47
48 long), compressed to 30% of original length (i.e., 120 ms), or compressed to 10% (i.e.,
49
50 40 ms). Congruency effects were found even for the strongly compressed primes. Mahr
51
52 and Wentura (2014) combined their Stroop task with a variation of perceptual load
53
54
55
56

57
58 ¹ Accordingly, the authors do not label their manipulation a “set size variation” but
59 “varied visual load”.
60

1
2
3
4 (Lavie, 1995, 2005). In detail, participants were presented with the colored target circle
5
6 in a ring-like arrangement along with seven other circles, either colored in non-target
7
8 colors (i.e., high perceptual load) or grey-shaded (i.e., low perceptual load). Stroop
9
10 effects were sharply increased in the high perceptual load condition. Note, the load
11
12 manipulation can be considered as almost equivalent to a visual search manipulation
13
14 with set size = 1 (low load) versus set size = 8 (high load). Thus, the increase of the
15
16 Stroop effect with larger set size can potentially be interpreted in terms of increased
17
18 search efficiency in the congruent priming condition.
19
20
21

22
23 However, Stroop effects are first of all interpreted in terms of response
24
25 facilitation/interference due to feature overlap between prime and target. Thus, it is not
26
27 straightforward to conclude that the experiment shows increased search efficiency (i.e.,
28
29 based on attentional processes). Therefore, with the present experiment we aimed to test
30
31 directly for the search efficiency hypothesis. Participants were presented with search
32
33 displays of set size 2, 8, or 16 (see Figure 1). In all trials, one of the stimuli was a target
34
35 color circle (i.e., a blue, green, yellow, or red circle) whereas distractor stimuli were
36
37 circles in non-target colors. The task was to classify a “gap” in the target circle as up or
38
39 down, thus, response-priming processes as underlying processes can be ruled out.
40
41
42

43
44 Auditory primes accompanied the visual presentation. Overall, we expected a priming
45
46 effect (i.e., faster responses to targets that were congruently compared to incongruently
47
48 primed). Most importantly, however, we expected that the search slope is less steep for
49
50 congruent trials compared to incongruent ones (i.e., indicating more efficient search for
51
52 the former one). To anticipate the discussion, such a result would imply that cross-
53
54 modal priming in visual search has an impact on attentional and working-memory
55
56 processes as suggested by recent theorizing: At a given point of time, working memory
57
58
59
60

1
2
3
4 may hold around four items (Cowan, 2000), for example, four target colors. However,
5
6 only a single item is considered having a privileged status as the focus of attention (e.g.,
7
8 Olivers, Peters, Houtkamp, & Roelfsema, 2011). Only this item is considered having the
9
10 potential to guide visual attention and hence is associated with increased search
11
12 efficiency. In line with these recent accounts, we assume that auditory primes might put
13
14 a corresponding working memory item into an active state within working memory, and
15
16 thereby into a prioritized status to bias attentional selection towards the semantically
17
18 matching visual item.
19
20
21

22 **Method**

23
24
25 The experiment was preregistered (<https://aspredicted.org/blind.php?x=hc26e5>).

26 **Participants**

27
28
29 Forty-three students (28 women, 15 men) from Saarland University took part in the
30
31 experiment in exchange for 4 Euro. The median age was 23 years (range 18 to 32
32
33 years). All had normal or corrected-to-normal vision. According to our preregistration,
34
35 the data of five further participants had to be discarded because of error rates that were
36
37 above three interquartile ranges above the third quartile with respect to the sample
38
39 distribution (*far out values*; Tukey, 1977).
40
41
42

43
44 For power calculations, we oriented ourselves on the difference between
45
46 facilitation effects for high versus low perceptual load found by Mahr and Wentura
47
48 (2014). In Experiment 1, the effect size of this difference was $d_z = .67$, in Experiment 2,
49
50 $d_z = .61$. Since we reduced the trial number (from 300 to 240) for the present
51
52 experiment and because we changed the task, we reduced the expected effect size to d_z
53
54 $= .50$. To detect an effect of $d_z = 0.50$ with power $1 - \beta = .95$ ($\alpha = .05$, one-tailed) a
55
56 sample size of $N = 45$ is needed. Because of a misunderstanding between authors and
57
58
59
60

1
2
3
4 experimenters with regard to the number of to-be-replaced participants (see above), the
5
6 recruitment phase ended with $N = 43$ valid persons. With $N=43$, power reduced to $1-\beta =$
7
8
9 .94.

10 11 **Design**

12
13 A 3 (congruency: congruent, incongruent, neutral) \times 3 (set size: 2, 8, 16 stimuli)
14
15 design will be employed with all factors manipulated within participants. Technically,
16
17 the congruency factor will be realized by a 5 (auditory prime: red, green, blue, yellow,
18
19 neutral) \times 4 (visual target: red, green, blue, yellow) design. That is, prime and target
20
21 features will be uncorrelated resulting in auditory priming without contingency (i.e.,
22
23 there is no benefit in expecting the primed color).
24
25
26

27 28 **Material**

29
30 Each target display contained two, eight, or sixteen visual stimuli presented in
31
32 random locations of a 4×4 matrix (18.5 x 18.5 cm, approx. $18.5 \times 18.5^\circ$ visual angle)
33
34 on a black background (see Figure 1). Each circle spanned 3.6° (diameter 73 pixels =
35
36 3.6 cm). One of these visual stimuli was the target and therefore appeared in one of the
37
38 four target colors (i.e., blue, green, yellow, red). The filler items were presented in
39
40 different non-target colors. These filler colors (e.g., pink, orange, turquoise) were
41
42 clearly distinguishable from the target colors. All circles had a gap in one of six
43
44 positions, three in the upper part of the stimulus ($-45^\circ, 0^\circ, 45^\circ$ with reference to the 12
45
46 o'clock position) and three in the lower part ($-45^\circ, 0^\circ, 45^\circ$ with reference to the 6
47
48 o'clock position).
49
50
51

52
53 We used the auditory word stimuli of Mahr and Wentura (2014), that is, time-
54
55 compressed sound files of exactly 120 ms length (i.e., compression to 30% of their
56
57 original length) of the one-syllable words red (“rot”, [ro:t]), green (“grün”, [gry:n]), blue
58
59
60

1
2
3
4 (“blau”, [blau]), and yellow (“gelb”, [gɛlp]). Four one-syllable non-words served as
5
6 neutral primes (“liez”, [li:ts], “tän”, [t ɛ:n], „nux“ [nʊks], and“ töff“ [toef]). The sounds
7
8 were presented over closed-ear headphones (AKGK511) and ranged in loudness from
9
10 68 to 72 dB SPL. The 30%-compression files were still understandable.

13 Procedure

14
15
16 Participants were individually seated in front of a 17-inch monitor (100 Hz
17
18 refresh rate, resolution 640×480 pixels) controlled by a personal computer in a sound
19
20 proof experimental cabin with dimmed light. Viewing distance was about 60 cm. The
21
22 experiment was conducted using E-prime software (E-prime 2.0). In each trial of the
23
24 experiment participants had to search for the target color and to categorize the location
25
26 of the target gap as up or down by pressing a corresponding key (keys f and j on a
27
28 standard keyboard; assignment counterbalanced). Participants were informed that the
29
30 auditory words would be time-compressed and not informative for the task.
31
32

33
34 To start each trial, participants pressed the space bar. In each trial, following a
35
36 1000 ms blank (black) screen, a white fixation cross appeared for 500 ms, followed by a
37
38 blank screen without sound for approx. 390ms (there was a slight jitter because of
39
40 loading the following audio file). The auditory prime was started during the presentation
41
42 of the blank screen, followed by the target screen with a SOA of 90 ms. The target
43
44 screen was presented until a response was given (see Figure 1).
45
46

47
48 Participants completed two practice phases. After being presented with all color
49
50 circles (i.e., the target and distractor circles), participants first completed 24 trials in
51
52 which they simply had to indicate by key-press whether a circle had a target color or
53
54 not. If necessary, this phase was repeated until participants reached at least 90%
55
56 accuracy. Then, they proceeded to a further practice block of 30 trials in which the main
57
58
59
60

experiment was practiced (i.e., a sub-sample of the full design was shown, with the different factors balanced). Feedback was given on each trial.

Afterwards, participants completed two experimental blocks of 120 trials each, composed of 24 neutral, 24 congruent, and 72 incongruent trials, randomly intermixed. The trial list was balanced with regard to the four colors (i.e., each of the 20 possible target-prime combinations was presented six times). Four warm-up-trials preceded each block (not included in the analyses). The experiment lasted for approximately 30 minutes.

Results

Unless otherwise noted, all effects referred to as statistically significant throughout the text are associated with p values below .05, two-tailed. We report all measures, manipulations, and exclusions in the study. Error rate was 8.5% ($SD = 10.1\%$). Trials with RTs below 150 ms or RTs greater than 1.5 interquartile ranges² above the third quartile with respect to the individual distribution of RTs were discarded (Tukey, 1977); this led to exclusion of 6.6% of trials. Mean RTs and error rates are reported in Table 1 (see also Figure 2).

Response Times. To test our hypotheses, we conducted a 3 (set size: 2 vs. 8 vs. 16) \times 3 (priming condition: neutral vs. congruent vs. incongruent) multivariate analysis of variance (MANOVA) for repeated measures (see O'Brien & Kaiser, 1985) with planned orthogonal contrasts. For the priming condition, the first contrast compared congruent versus incongruent trials (i.e., whether priming of auditory primes onto the

² In the preregistration, we specified that we would discard RTs that were 1.5 interquartile ranges above the third quartile or below the first quartile (Tukey, 1977). There are no such outliers at the lower end of the individual distributions. However, we removed RTs below 150 ms (i.e., preparatory responses; 0.2% of all trials) in line with usual and meaningful practice (which we forgot to mention in the preregistration). Inclusion or exclusion of these very few trials made no essential difference in results.

1
2
3
4 visual search targets occurred); it is the contrast of main interest (i.e., the overall
5
6 priming effect). Accordingly, for the second contrast, reaction times of congruent and
7
8 incongruent trials will be averaged and contrasted with those for the neutral stimuli.
9
10 This contrast was of minor interest in the present study. The set size factor was
11
12 transformed in linear and quadratic trend.
13
14

15
16 As expected, the analysis revealed a significant main effect of congruency,
17
18 $F(2,41) = 12.58, p < .001, \eta_p^2 = .725$. The a priori contrast of main interest (congruent
19
20 vs. incongruent trials) revealed a significant difference, $F(1,42) = 25.71, p < .001, \eta_p^2 =$
21
22 $.380$; responses in congruent trials were faster than responses in incongruent trials (see
23
24 Table 1 and Figure 2). The second contrast, comparing neutral with relevant prime
25
26 words (congruent and incongruent pooled), was not significant, $F(1,42) = 3.68, p =$
27
28 $.062, \eta_p^2 = .081$.
29
30
31

32
33 The main effect for set size was significant as well, $F(2,41) = 124.40, p < .001,$
34
35 $\eta_p^2 = .859$. This effect was dominantly due to a linear trend, $F(1,42) = 214.71, p < .001,$
36
37 $\eta_p^2 = .836$, with slower responses for larger displays. The quadratic trend was
38
39 significant as well, $F(1,42) = 16.57, p < .001, \eta_p^2 = .283$, indicating that RTs in the
40
41 medium set size condition are not exactly the average of the RTs of small and large
42
43 display-size (see Figure 2).
44
45

46
47 Most importantly, there was a significant interaction of priming and set size,
48
49 $F(4,39) = 5.89, p < .001, \eta_p^2 = .377$. This interaction is exclusively due to the
50
51 interaction of the congruent versus incongruent contrast of the priming factor and the
52
53 linear trend of set size, $F(1,42) = 12.58, p < .001, \eta_p^2 = .230$ ($F(1,42) = 1.74, p = .195,$
54
55 $\eta_p^2 = .040$ for the neutral vs. congruent/incongruent contrast \times linear trend; $F_s < 1$ for
56
57 the two remaining combinations). As can be seen in Figure 2, the effect of set size is
58
59
60

1
2
3
4 reduced in congruent trials compared to incongruent trials providing evidence for our
5
6 search efficiency hypothesis.
7

8
9 Although all priming effects for the three set sizes were significant, they
10
11 increased in terms of mean RT difference and effect size with increasing set size: $t(42)$
12
13 $= 2.53, p = .015, d_z = 0.39$ for set size = 2; $t(42) = 4.06, p < .001, d_z = 0.62$ for set size =
14
15 8; $t(42) = 5.04, p < .001, d_z = 0.77$ for set size = 16. Additionally, Table 1 includes costs
16
17 and benefits; we will discuss them in the *Discussion* section.
18
19

20
21 **Error Rates.** A 3 (set size) \times 3 (priming condition) MANOVA for repeated
22
23 measures revealed no significant overall effects, $F(2,41) = 2.71, p = .078, \eta_p^2 = .117$ for
24
25 priming, $F < 1$ for set size, and $F(4,39) = 1.18, p = .336, \eta_p^2 = .108$ for the interaction.
26
27 There were more errors in the incongruent condition compared to the congruent one,
28
29 $F(1,42) = 4.10, p = .049, \eta_p^2 = .089$, mirroring the RT priming effect. Numerically the
30
31 priming effects increase as well from low set size to large set size; the corresponding
32
33 interaction contrast was, however, not significant, $F(1,42) = 1.69, p = .201, \eta_p^2 = .039$
34
35 ($F < 1$ for all remaining contrasts). In sum, there are no indications of a speed-accuracy
36
37 tradeoff with regard to the RT effects.
38
39
40

41
42 **Linear Mixed Model analyses.** To corroborate the traditional analysis, we
43
44 analyzed RTs (of correct responses; outliers discarded; see above) as a function of
45
46 priming condition (coded +1 congruent, -1 incongruent), the set size (i.e., 2 / 8 /16), and
47
48 their interaction by using linear mixed model analysis (lmerTest package, Kuznetsova,
49
50 Brockhoff, & Christensen, 2016, Bates, Maechler, Bolker, & Walker, 2015; R-Core-
51
52 Team, 2016). We allowed random intercepts and slopes for participants. Set size had a
53
54 weight of $b = 18.9$ ms (SE = 1.0 ms), $t(44.35) = 18.46, p < .001$, indicating that with
55
56 each additional element in the display, search RTs increased by 18.9 ms. The interaction
57
58
59
60

of set size with priming was significant as well, $b = -2.5$ ms (SE = 0.7 ms), $t(122.75) = 3.40$, $p < .001$. Slopes were less steep for congruent ($b = 16.4$ ms) and steeper for incongruent pairs ($b = 21.4$ ms).³

Discussion

Our study provides evidence for increased search efficiency as a mechanism in cross-modal visual search, indicating that attentional and working memory processes are altered by a congruent prime. Specifically, in a visual search experiment, we found faster target processing if the semantics of a brief auditory prime matched the target category. Importantly, this congruency effect increased with increased set size, or – in other words – the search slope was significantly flatter in case of congruent priming compared to incongruent priming. This is an important result because it relates cross-modal priming to recent research in working memory research. We thereby show that representations can be cross-modally activated, contained in an active state in working memory, and guide visual attention.

Since there are only four possible target colors that are repeated over and over, it is plausible to assume that congruent priming does not facilitate retrieval of an item from long-term memory (as is assumed in typical semantic priming studies; see McNamara, 2005, 2013, for reviews). Rather, it is more plausible to assume that the four relevant target items are actively maintained in working memory throughout the experiment.

Thus, facilitatory priming by an auditory prime can be explained as bringing the primed

³ A more complex analysis including the neutral trials (with two contrast codes for priming and therefore two interaction terms) yielded almost the same result for the interaction of contrast congruent versus incongruent with set size, $b = -2.5$ ms (SE = 0.7 ms), $t(101.17) = 3.35$, $p = .001$, and a non-significant result for the interaction of contrast neutral versus congruent/incongruent with set size, $b = -1.0$ ms (SE = 1.0 ms), $t(81.50) = 1.00$, $p = .323$.

1
2
3
4 working memory item – metaphorically speaking – into the “foreground” (Mahr &
5
6 Wentura, 2014).

7
8
9 Thus, our study adds to recent theorizing in working memory research (Garavan,
10
11 1998; McElree, 2006; Nee & Jonides, 2011; Oberauer, 2002; Olivers et al., 2011).
12
13 Although working memory may hold around four items, only a single item is assumed
14
15 to have a prioritized status and thus the potential to guide visual attention by increasing
16
17 visual sensitivity to the corresponding visual stimulus (Desimone & Duncan, 1995;
18
19 Olivers, Meijer, & Theeuwes, 2006). The prioritized working memory item can thus be
20
21 regarded as an active template that directly resonates with corresponding visual input
22
23 (Olivers et al., 2011).
24
25

26
27 The results of the neutral priming condition (in comparison to the congruent and
28
29 incongruent ones) give further insights into the underlying mechanisms: For the small
30
31 set size, incongruent and neutral conditions did not differ; that is, there were no costs of
32
33 incongruent priming. In contrast, for the largest set size, cost and benefits were more
34
35 balanced. This is plausible as, in the small set size displays with almost no competition
36
37 of distractors, the target is immediately given. In the congruent condition, the target
38
39 matches the active working memory item; processing the item according to the task
40
41 starts immediately. In the incongruent condition, the target can be considered a visual
42
43 prime that first alters the state of the working memory items – the target-matching item
44
45 switches now from the passive to the active state; hence, start of the task-relevant
46
47 process is a bit prolonged. The same happens in the neutral prime condition, if we
48
49 assume that none of the four working memory items is prioritized in the neutral
50
51 condition.
52
53
54
55
56
57
58
59
60

1
2
3
4 In the large set size display, by contrast, the active working memory item guides
5 visual search. This has beneficial effects in the congruent case in terms of search
6 efficiency. However, it also seems to have detrimental effects for those working
7 memory items that are currently in the passive state, as we found also costs associated
8 with incongruent trials. It is up to further research to explore whether there is a genuine
9 inhibition for the passive items or whether the detrimental effect is a by-product of
10 attentional effects caused by the active item (e.g., in the present context it might lead to
11 attention allocation on distractors that come closest to the color denoted by the active
12 item).

13
14
15
16
17
18
19
20
21
22
23
24
25 To conclude, our study provided a direct test of the search efficiency hypothesis,
26 showing that cross-modal priming in a visual search task impacts attentional and
27 working memory processes as indicated by flatter search slopes for congruent compared
28 to incongruent trials.
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

Open Practices Statement

The data and materials for the experiment are available at

(https://osf.io/uxm4r/?view_only=f56790de0f9b45bbbca28e0d7f7008ae); the

experiment was preregistered (<https://aspredicted.org/blind.php?x=hc26e5>).

For Review Only

References

- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1-48.
doi:10.18637/jss.v067.i01
- Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87-185.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193-222.
- Garavan, H. (1998). Serial attention within working memory. *Memory & Cognition*, *26*, 263-276. doi:10.3758/bf03201138
- Iordanescu, L., Grabowecky, M., Franconeri, S., Theeuwes, J., & Suzuki, S. (2010). Characteristic sounds make you look at target objects more quickly. *Attention, Perception, & Psychophysics*, *72*, 1736-1741. doi:10.3758/APP.72.7.1736
- Iordanescu, L., Grabowecky, M., & Suzuki, S. (2011). Object-based auditory facilitation of visual search for pictures and words with frequent and rare targets. *Acta Psychologica*, *137*, 252-259. doi:10.1016/j.actpsy.2010.07.017
- Iordanescu, L., Guzman-Martinez, E., Grabowecky, M., & Suzuki, S. (2008). Characteristic sounds facilitate visual search. *Psychonomic Bulletin & Review*, *15*, 548-554. doi:10.3758/PBR.15.3.548
- Knoeferle, K. M., Knoeferle, P., Velasco, C., & Spence, C. (2016). Multisensory brand search: How the meaning of sounds guides consumers' visual attention. *Journal of Experimental Psychology: Applied*, *22*, 196-210. doi:10.1037/xap0000084

- 1
2
3
4 Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). lmerTest: Tests in
5
6 linear mixed effects models. R package version 2.0-32. [https://CRAN.R-](https://CRAN.R-project.org/package=lmerTest)
7
8 [project.org/package=lmerTest](https://CRAN.R-project.org/package=lmerTest).
9
10
11 Lavie, N. (1995). Perceptual load as a necessary condition for selective attention.
12
13 *Journal of Experimental Psychology: Human Perception and Performance*, *21*,
14
15 451-468.
16
17
18 Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in*
19
20 *Cognitive Sciences*, *9*, 75-82.
21
22
23 Mahr, A., & Wentura, D. (2014). Time-compressed spoken word primes crossmodally
24
25 enhance processing of semantically congruent visual targets. *Attention,*
26
27 *Perception, & Psychophysics*, *76*, 575-590. doi:10.3758/s13414-013-0569-z
28
29
30 Mahr, A., & Wentura, D. (2018). Does a “stoplight!” improve processing a stoplight?
31
32 Cross-modal influences of time-compressed spoken denotations on automotive
33
34 icon classification. *Journal of Experimental Psychology: Applied*, *24*, 92-107.
35
36
37 McElree, B. (2006). Accessing recent events. In B. H. Ross (Ed.), *Psychology of*
38
39 *Learning and Motivation: Advances in Research and Theory, Vol 46* (Vol. 46,
40
41 pp. 155-200). San Diego: Elsevier Academic Press Inc.
42
43
44 McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word*
45
46 *recognition*. New York: Psychology Press.
47
48
49 McNamara, T. P. (2013). Semantic memory and priming. In A. F. Healy & R. W.
50
51 Proctor (Eds.), *Handbook of psychology, Vol. 4: Experimental psychology (2nd*
52
53 *ed.)*. (pp. 449-471). Hoboken, NJ: John Wiley & Sons Inc.
54
55
56
57
58
59
60

- 1
2
3
4 Nee, D. E., & Jonides, J. (2011). Dissociable contributions of prefrontal cortex and the
5
6 hippocampus to short-term memory: Evidence for a 3-state model of memory.
7
8 *Neuroimage*, 54, 1540-1548. doi:10.1016/j.neuroimage.2010.09.002
9
10
11 Oberauer, K. (2002). Access to information in working memory: Exploring the focus of
12
13 attention. *Journal of Experimental Psychology: Learning, Memory, and*
14
15 *Cognition*, 28, 411-421.
16
17
18 Olivers, C. N. L., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven
19
20 attentional capture: Visual working memory content affects visual attention.
21
22 *Journal of Experimental Psychology: Human Perception and Performance*, 32,
23
24 1243-1265.
25
26
27 Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states
28
29 in visual working memory: When it guides attention and when it does not.
30
31 *Trends in Cognitive Sciences*, 15, 327-334.
32
33
34 R-Core-Team. (2016). R: A language and environment for statistical computing. R
35
36 Foundation for Statistical Computing, Vienna, Austria. [https://www.R-](https://www.R-project.org/)
37
38 [project.org/](https://www.R-project.org/)
39
40
41 Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.
42
43
44 Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention* (pp. 14-73). Hove,
45
46 UK: Psychological Press.
47
48 Zweig, L. J., Suzuki, S., & Grabowecky, M. (2015). Learned face–voice pairings
49
50 facilitate visual search. *Psychonomic Bulletin & Review*, 22, 429-436.
51
52 doi:10.3758/s13423-014-0685-3
53
54
55
56
57
58
59
60

Author Note

Dirk Wentura, Michaela Rohr, and Timea Folyi, Department of Psychology,
Saarland University, Saarbrücken, Germany.

The research reported in this article was supported by a grant from the Deutsche
Forschungsgemeinschaft (DFG) to Dirk Wentura (WE 2284/9-1). Correspondence
concerning this article should be addressed to Dirk Wentura, Department of
Psychology, Saarland University, Campus A2 4, D-66123 Saarbrücken, Germany, e-
mail: wentura@mx.uni-saarland.de.

For Review Only

Table 1

Mean RTs (in ms) and Accuracy (in %) as a function of set size and semantic congruency conditions; priming effects (PE), costs, and benefits (SE in parentheses)

Set Size	Semantic congruency condition			PE	Benefit	Cost
	Neutral	Congr.	Incongr.			
<i>RT</i>						
2	768	726	754	27 (11)	42 (15)	-14 (13)
8	922	865	937	72 (18)	57 (17)	15 (15)
16	1009	959	1054	95 (19)	50 (19)	45 (14)
<i>Accuracy</i>						
2	8.0	7.8	9.0	1.1 (1.3)	0.1 (1.7)	1.0 (1.3)
8	7.0	7.3	8.9	1.7 (1.6)	-0.3 (1.5)	1.9 (1.0)
16	8.7	6.4	9.7	3.3 (1.3)	2.3 (1.3)	1.0 (1.2)

Note. PE = $RT_{\text{incongruent}} - RT_{\text{congruent}}$; Benefit = $RT_{\text{neutral}} - RT_{\text{congruent}}$; Cost = $RT_{\text{incongruent}} - RT_{\text{neutral}}$

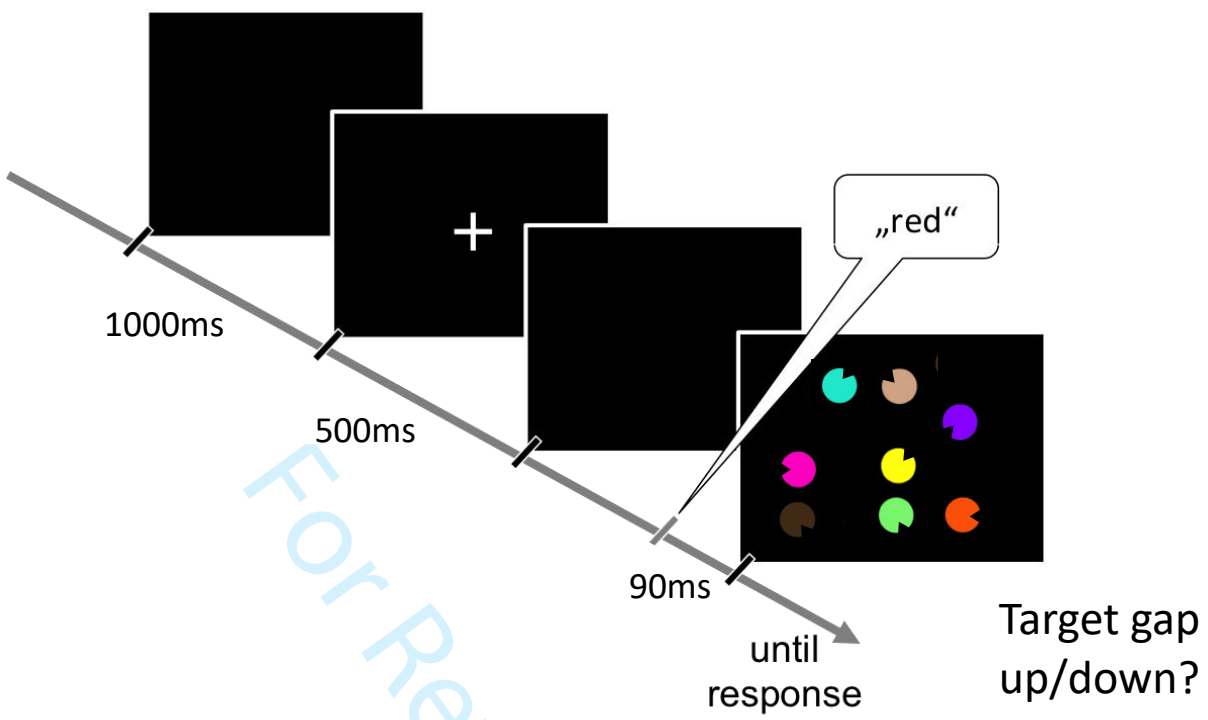
Figure Captions*Figure 1*

Trial sequence (not drawn to scale; example: incongruent trial; set size 8). The prime word (here: red) is presented via headphones 90 ms prior to the visual target (here: yellow). The correct response is here a right key press (because the yellow circle has its gap in the upper part).

Figure 2

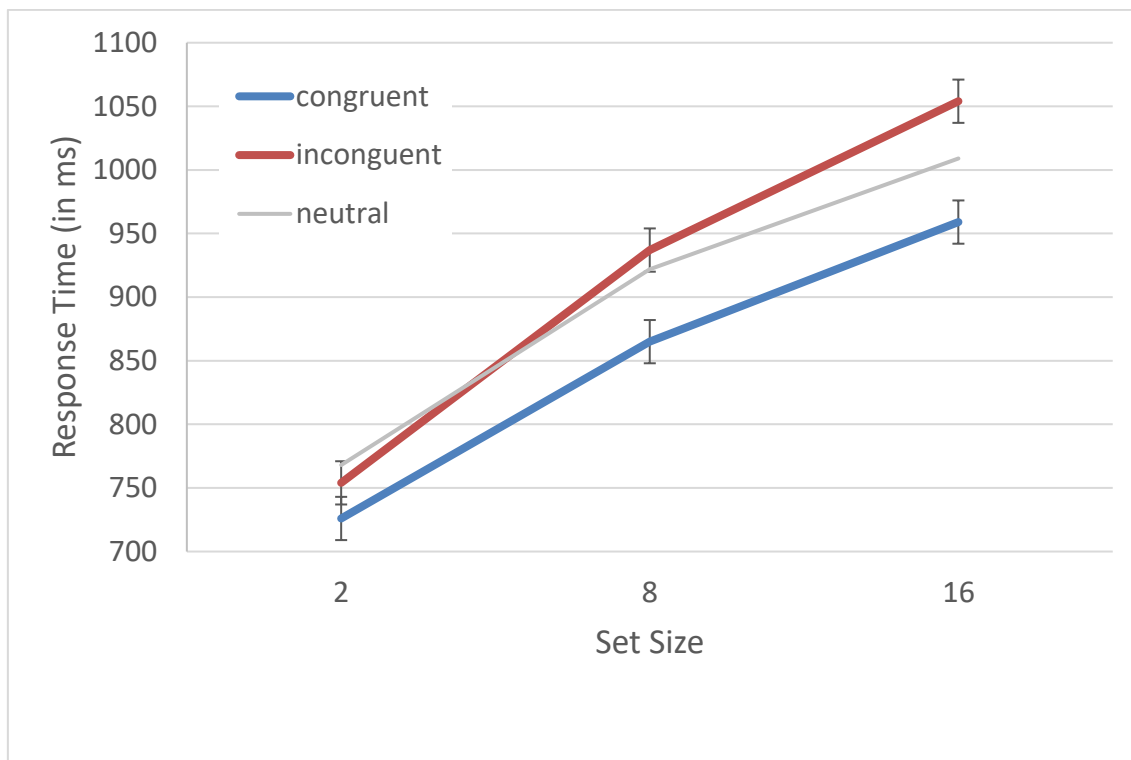
Mean RTs as a function of set size and type of priming. Error bars are 95% within-subject confidence intervals (Jarmasz & Hollands, 2009) for the 2 (priming condition: congruent vs. incongruent) \times 3 (set size) interaction effect.

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



For Review Only

Cross-modal enhancement of search efficacy 23



review Only