Journal of Cognitive Psychology

ISSN: 2044-5911 (Print) 2044-592X (Online) Journal homepage: https://www.tandfonline.com/loi/pecp21

Effects of visual complexity and ambiguity of verbal instructions on target identification

Tim van Marlen, Margot van Wermeskerken & Tamara van Gog

To cite this article: Tim van Marlen, Margot van Wermeskerken & Tamara van Gog (2019) Effects of visual complexity and ambiguity of verbal instructions on target identification, Journal of Cognitive Psychology, 31:2, 206-214, DOI: 10.1080/20445911.2018.1552700

To link to this article: https://doi.org/10.1080/20445911.2018.1552700

Published online: 29 Nov 2018.

JOURNAL OF COGNITIVE PSYCHOLOGY
2019, VOL. 31, NO. 2, 206–214
https://doi.org/10.1080/20445911.2018.1552700

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Effects of visual complexity and ambiguity of verbal instructions on target identification

Tim van Marlen, Margot van Wermeskerken and Tamara van Gog

Department of Education, Utrecht University Utrecht, Netherlands

ABSTRACT
Research has shown that visual complexity and the ambiguity of verbal information affect the speed and accuracy of locating targets during visual search. The higher the visual complexity and description ambiguity, the slower and poorer the target identification performance. Because these factors are seldom studied in combination (even though they regularly co-occur), it is unclear whether they would interact. Therefore, in two experiments, participants viewed images that displayed cartoon-like characters and had to correctly identify a character from a verbal description under conditions of low/high visual complexity and low/high description ambiguity (manipulated within-subjects). Results revealed that high ambiguity descriptions resulted in lower accuracy and slower response times. However, our manipulation of visual complexity did not affect performance or response times either in itself or in interaction with verbal ambiguity. Findings are discussed in terms of theoretical and practical implications, for instance, for multimedia learning.

ARTICLE HISTORY
Received 6 April 2017
Accepted 21 November 2018

KEYWORDS
Visual complexity; ambiguity resolution; visual search; visual world paradigm

Research on visual search has been concerned with the question of how the visual complexity of the stimuli (Wolfe, 1994) or the ambiguity of the (verbal) task instructions (Louwerse & Bangerter, 2010) affect the speed and accuracy of locating a target. However, these factors (i.e. visual complexity and verbal ambiguity) have primarily been studied in isolation, so whether and how they interact to influence visual search is unknown. The potential interaction of visual complexity and verbal ambiguity on visual search is important as it determines the specificity of the target template (i.e. the representation of the features of the target). A clear target template is not always readily available prior to commencing the search. For instance, search can be less effective if the verbal target description is ambiguous, leaving too many potential targets candidates, and the visual display is too complex to effortlessly spot the target.

The present study was designed to test the prediction that both high visual complexity and high description ambiguity would result in poorer target identification performance (lower accuracy and slower response times).

Visual complexity and specificity

The visual complexity of an image or task display can be an important characteristic in determining the amount of visual search required to locate and identify a target. The concept of visual complexity has many definitions, discussion of which goes beyond the scope of this article (for reviews see Donderi, 2006b; Forsythe, 2009). Although different definitions exist, most agree with the basic definition that visual complexity refers to the amount of detail present within an image. By visual complexity, we mean the number of visual features present, the amount of colour differences, and contrast differences. Besides this subjective definition, visual complexity can be quantified by looking at the file size of a compressed digital image (Chikhman, Bondarko, Danilova, Goluzina, & Shelepin, 2012; Donderi, 2006a; Marin & Leder, 2016). That is, the more visually complex an uncompressed image is, the more

CONTACT Tim van Marlen t.v.a.vanmarlen@uu.nl

difficulty the compression algorithm has compressing the image, thus resulting in a larger image file and vice versa. Research has shown that visual search becomes less efficient (Wolfe, 1994) and less accurate (Davis, Shikano, Peterson, & Michel, 2003; Neider & Zelinksky, 2011; Wolfe, Oliva, Horowitz, Butcher, & Bompas, 2002) with increasing visual complexity. For instance, Neider and Zelinksky (2011) demonstrated that when visual complexity increases (i.e. clutteredness of an image), individuals became less efficient in searching for a target building embedded in a scene image. However, in such studies, the target is cued beforehand by a specific target image providing the individual with a specific target template to aid visual search, which is not always readily available in everyday life. Therefore, how visual complexity affects visual search might also depend on the specificity of the target template and this specificity might in turn also be affected by the verbal ambiguity.

Indeed, there are many studies indicating that the specificity of a target template can influence the efficiency of visual search (Castelhano & Heaven, 2010; Hout & Goldinger, 2015; Malcolm & Henderson, 2009; Schmidt & Zelinsky, 2009; Yang & Zelinsky, 2009). In most of these studies, the target template is either a picture of the target or the target object name or category. Generally, it is found that visual search is most efficient when the target is cued with a picture (e.g. picture of a mug) compared to when it is cued with an object name (e.g. the text "mug"), because an object name is less specific about the target than the picture (Castelhano & Heaven, 2010; Malcolm & Henderson, 2009; Schmidt & Zelinsky, 2009). However, in most of these studies, the visual complexity of the images that are to be searched for the target is either low (e.g. depicting objects with little visual details) or is not manipulated. An exception is the study of Godwin and colleagues (Godwin, Walenchok, Houpt, Hout, & Goldinger, 2015). In this study participants had to classify objects in a dual target search task (i.e. classifying two targets simultaneously) and were given a picture cue of the target before each block of trials. The targets were either simple images (i.e. Landolt C's) or complex images (i.e. teddy bears and butterflies). Contrary to what one might expect, it was found that participants were faster in rejecting the distractors in the case of high complexity as compared to low complexity images, possibly due to the larger number of features present in the complex images that could be used to identify the objects.

In sum, research indicates that visual search can be affected by the visual complexity of the stimuli and that target cue images facilitate visual search (due to the higher specificity of the target template) compared to object names as target cue search templates. Although some studies mentioned above-used target templates cues consisting of object or category names, these textual target cues were relatively unambiguous. In the next section, we discuss how ambiguously perceived verbal task descriptions can affect visual attention.

Verbal ambiguity

Besides the visual complexity of a task, the verbal ambiguity of a task description might also influence the amount of visual search required within a learning task. Support for the idea that the ambiguity of verbal descriptions influences the way people attend visual information comes from research using the visual world paradigm (Allopenna, Magnuson, & Tanenhaus, 1998; Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995; for a review see Huettig, Rommers, & Meyer, 2011). In this paradigm, people see an image depicting several objects while hearing a verbal description that is ambiguous about some of the objects. For instance, in the study by Allopenna et al. (1998) people saw an image displaying a beaker, a beetle, a speaker, and a carriage and heard the word *beaker*. Eye tracking data revealed that just after the onset of the word "beaker" the probability of fixating the image of the beaker did not differ from fixating the images of beetle and speaker, but once the whole word "*beaker*" has been heard and the ambiguity is resolved the probability of fixating the beaker is higher than the probability of fixating the other objects. This finding suggests that the ambiguity of verbal information (in this case arising from the visual context) affects target identification.

Comparable results were found in a study by Louwerse and Bangerter (2010) using an adapted version of the visual world paradigm in which participants were shown images containing a grid of 3 × 4 cartoon faces accompanied by a verbal description about one of the cartoon faces. Half of the verbal descriptions contained an additional ambiguous location description (i.e. *John is in the middle* …), whereas the other half did not. In addition, in half of the trials there was an ambiguous gesture cue (e.g. a pointing hand) pointing at the correct row, while the other half did not. Note that the presence of both types of cues resolves the ambiguity regarding the target location. Both verbal cues and gesture cues guided attention (i.e. were better than no descriptions/cues for target identification), but their combination most improved target identification, as indicated by the higher percentage of fixations on the correct cartoon face for the conditions with the extra location cue(s). In this study, the ambiguity arose from the lack of specificity of the verbal description (which could be resolved by cues in the visual context).

Present study

The literature discussed above seems to indicate that both conditions of high visual complexity and conditions of high verbal ambiguity negatively affect the accuracy and speed of target identification. Yet, to the best of our knowledge, the effects of visual complexity and verbal description ambiguity have not yet been investigated in a single study, which renders it difficult to draw conclusions on these factors' individual influence on visual search processes and their possible interaction. Therefore, we used a variation of the visual world paradigm (cf. Louwerse & Bangerter, 2010), to explore the effects of visual complexity and verbal description ambiguity on target identification performance and response times. Participants saw images depicting 12 cartoon characters on a 4 × 3 grid while hearing a verbal description of one of the characters, which they had to identify. Visual complexity (low vs. high) was defined as the number of visual features contained within a display. Verbal description ambiguity (low vs. high) was defined as the number of characters referred to by the first two cues presented in the verbal description. We predicted that both high visual complexity and high verbal description ambiguity would result in lower accuracy and slower response times on accurate trials than low visual complexity and low verbal description ambiguity. An open question is whether the effects would be additive, thus even more pronounced in the combined condition with high visual complexity and high verbal description ambiguity. To investigate this question two experiments were conducted: Experiment 1 was conducted online with Amazon's Mechanical Turk and Experiment 2 was conducted in the lab.

Experiment 1

Method

Participants and design

Participants were 250 U.S.A. citizens recruited online with Amazon's Mechanical Turk and received $0.80 payment for participation. Ten participants were excluded due to incomplete data, resulting in 239 participants ($M_{age} = 38.14$, $SD_{age} = 12.44$; 146 females). A 2 × 2 factorial within-subjects design was used, with visual complexity (high vs. low) and verbal ambiguity (high vs. low) as factors.

Materials

The experiment was run in Qualtrics software (www.qualtrics.com).

Target displays. Cartoon characters were created with an online avatar creator tool (www.sp-studio.de). The rectangular area covering a cartoon character was 230 × 230 pixels in size and the total image consisting of 4 × 3 cartoon characters was 1280 × 920 pixels. Thus, each display for a target identification trial featured 12 cartoon characters (six males/six females, with/without glasses, with white/olive/dark-coloured skin, and with blond/brown/black/red hair). The *low* and *high* visual complexity conditions differed in the number of variations amongst characters. The low-complexity condition featured characters with one female and one male outfit, one hair style, and two hair colours. The high-complexity condition featured characters with two female and two male outfits (with more visual details than the low-complexity condition), two hair styles, two hair colours, who additionally carried accessories (i.e. headphones, food, bottles). Above each character, a number was shown (i.e. 1–12; see Figure 1). A play button for the audio was located underneath each image (this was done automatically by Qualtrics).

To verify with an objective measure[1] that the images used in the different conditions indeed

[1]Upon request of a reviewer we additionally collected subjective complexity ratings. Thirty-four participants ($M_{age} = 25.00$, $SD_{age} = 4.53$) who did not participate in Experiment 1 or 2, were presented with the low and high visual complexity images one-by-one (in random order), and asked to rate on a 7-point scale (with 1 = 'very simple' and 7 = 'very complex') how visually complex they perceived the images to be. A paired *t*-test on the average ratings of the low ($M = 2.93$, $SD = 0.95$) vs. high ($M = 4.22$, $SD = 1.04$) complexity images revealed a large effect, $t(33) = 7.40$, $p < .001$,

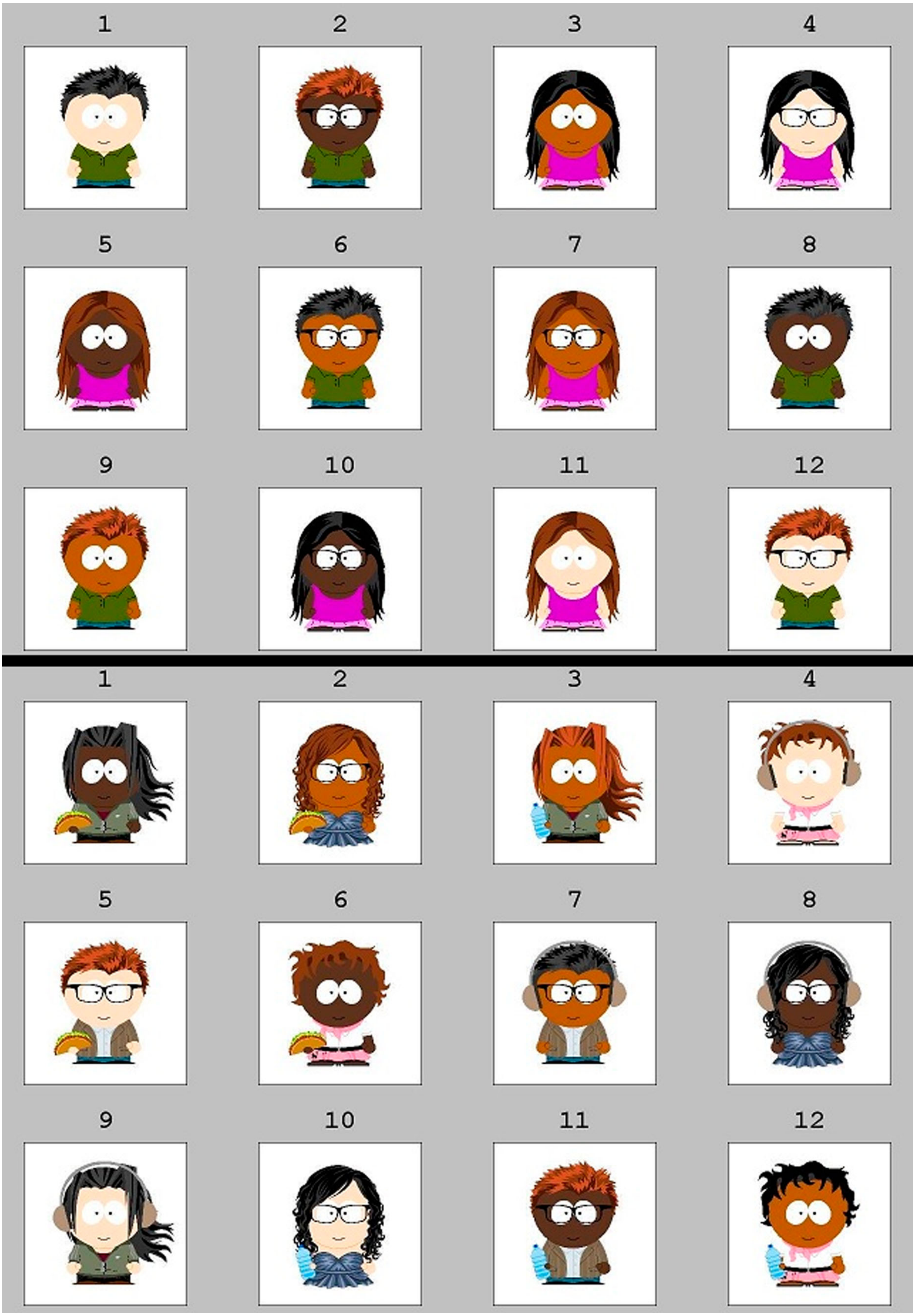

Figure 1. An example image displaying the cartoon characters of the low visual complexity condition (top) and the high visual complexity condition (bottom).

differed in visual complexity, a paired *t*-test was conducted with the file size of the JPEG images as the dependent variable (cf. Chikhman et al., 2012; Donderi, 2006a; Marin & Leder, 2016). The average file size of the images in the low-complexity condition ($M = 106.42$ KB, $SD = 0.67$) was indeed lower than that of the images in the high-complexity condition ($M = 117.17$ KB, $SD = 0.94$), t (22) = 32.34, $p < .001$, $r = .99$.

Twenty-five displays were created: one practice, 12 low visual complexity and 12 high visual complexity displays. The 24 low/high complexity displays were used twice: once with a low verbal ambiguity description and once with a high verbal ambiguity description (i.e. 48 trials: 12 low complexity/low ambiguity, 12 low complexity/high ambiguity, 12 high complexity/low ambiguity, and 12 high complexity/high ambiguity).

For each trial, one of the twelve cartoon characters had to be identified (i.e. target). Within each of the four conditions (high/low complexity x high/low ambiguity), each character number position (1–12) was used as a target location only once.

Verbal descriptions. The verbal descriptions were spoken by a female voice and lasted on average 7.62 s ($SD = 0.43$). Each verbal description (e.g. "A person with brown hair, next to a person without glasses, has an olive colored skin.") contained three cues necessary to correctly identify the target character (e.g. hair colour, skin colour). The verbal descriptions were chosen in such a way that the number of cartoon characters that fit the description decreased along with each of the three cues of the description. In the low-ambiguity condition, after the first, second, and third cue, there were four, two, and one possible cartoon characters left who fit the description, respectively. In the high-ambiguity condition, there were six, four, and one possible cartoon characters left who fit the description after the first, second, and, third cue, respectively.

Procedure

After participants agreed to take part in the study on Amazon's Mechanical Turk, they were redirected to the Qualtrics online survey. Participants were instructed that they would see displays featuring 12 cartoon characters, while hearing descriptions of one of the characters, and that their task was to correctly identify the described character and type in the correct character number as fast as possible. After a practice trial, participants received the 48 trials (12 per condition) in random order. After completion of the task, participants provided demographic data. The experiment lasted approximately 20 min.

Data analysis

Within each condition, target identification accuracy was measured as the proportion of correctly identified targets. Average response times per condition on *correct* trials were calculated. 21 participants were excluded from analyses due to low performance on the task (i.e. > 3*SD* below average; $n = 2$) or unfeasible display resolution (i.e. participants had to scroll in order to *view* the full stimulus display, $n = 19$). Moreover, many participants reported that the cue "a person with a *light* colored skin" was often mistaken for the cue "a person with a *white* colored skin". To explore whether this confusion had affected the performance of the participants on these trials, a paired *t*-test on proportion correct was conducted between trials that contained the cue "a person with a light colored skin" and the remaining trials. This analysis confirmed that performance was lower on trials with the "light" coloured skin cue ($M = .53$, $SE = .02$) as compared to the other trials ($M = .82$, $SE = .01$), $t(217) = -15.66$, $p < .001$, $r = .73$. Therefore, trials containing this confusing cue ($n = 16$) were excluded from further analyses, leaving eight trials in each condition. Regarding the response times analyses, 33 additional participants were excluded from the analyses due to very long response times (i.e. > 3*SD* above average; $n = 4$), RT data not being logged ($n = 1$), unfeasible display resolution (i.e. participants had to scroll down to be able to *type* in their answer; $n = 26$) or having no correct responses within a condition ($n = 2$).

Results and discussion

Data (Table 1) were analysed with 2 (visual complexity: low/high) x 2 (verbal ambiguity: low/high) repeated measures ANOVAs. On target identification accuracy, in contrast to our prediction this analysis revealed no main effect of visual complexity, F (1, 217) = 1.75, $p < .188$, $\eta_p^2 = .01$. However, in line with our prediction there was a main effect of verbal ambiguity, F (1, 217) = 15.77, $p < .001$, $\eta_p^2 = .07$,

$r = .78$. Thus, both according to an objective (file size) and subjective measure, there was a difference in visual complexity between low and high complexity images.

Table 1. Mean (and SD) proportion correct identifications and response times in Experiment 1 and in Experiment 2.

	Low visual complexity		High visual complexity	
	Low verbal ambiguity	High verbal ambiguity	Low verbal ambiguity	High verbal ambiguity
Experiment 1 ($N = 239$)				
Proportion correct ($N = 218$)	0.85 (0.21)	0.80 (0.22)	0.82 (0.20)	0.80 (0.21)
Response time (s) ($N = 185$)	17.60 (5.35)	20.95 (6.87)	20.82 (9.13)	21.76 (7.75)
Experiment 2 ($N = 34$)				
Proportion correct ($N = 33$)	0.83 (0.21)	0.74 (0.18)	0.85 (0.16)	0.73 (0.16)
Response time (s) ($N = 29$)	14.66 (2.38)	16.39 (2.60)	15.06 (2.04)	16.81 (4.38)

indicating that higher ambiguity resulted in less accurate performance, but only on the low-complexity trials as indicated by an interaction effect, F (1, 217) = 4.76, $p < .030$, $\eta_p^2 = .02$, with low-complexity/low-ambiguity leading to the highest accuracy, t (217) = 2.46, $p = .015$, $r = .16$.

In line with our predictions, the analyses regarding identification speed revealed main effects of visual complexity, F (1, 184) = 26.54, $p < .001$, $\eta_p^2 = .13$, and verbal description ambiguity, F (1, 184) = 28.57, $p < .001$, $\eta_p^2 = .13$, with high-complexity and high ambiguity resulting in slower response times. However, this effect was again qualified by an interaction effect, F (1, 184) = 8.31, $p = .004$, $\eta_p^2 = .04$, indicating that verbal ambiguity only affected response times on low-complexity trials, t (184) = 5.50, $p < .001$, $r = .38$, with low-complexity/low-ambiguity leading to the fastest response times.

In sum, the results of Experiment 1 revealed that participants were slower and less accurate in identifying the correct cartoon character for the high ambiguity trials when the visual complexity was low. Although the results are in line with previous research stating that the more specific the target template is (i.e. in this case the low ambiguity conditions) the better the performance, we did not find the expected interaction in which performance would deteriorate even further in high ambiguity trials that were also high in complexity. However, the results have to be interpreted with caution, as participants were able to replay the verbal descriptions. This could have affected the results on both speed and accuracy (e.g. the lack of effect of verbal ambiguity on high-complexity trials might be due to participants replaying the description). To ensure that participants would only hear the verbal description once, a second experiment was conducted in the lab.

Experiment 2

The set-up and materials in Experiment 2 were the same as in Experiment 1, except for three adjustments: First, participants were tested in a controlled lab environment instead of online. Second, participants were explicitly instructed to listen to the verbal descriptions only once (which was controlled for by the experimenter). Third, due to the confusion caused by the cue concerning skin colour, the formulation of this cue was changed. With these changes the aim of Experiment 2 was similar to the aim of Experiment 1, but with the difference that the verbal description was presented only once. We predicted to see similar effects of visual complexity and task ambiguity as seen in Experiment 1 but with somewhat lower overall performance due to the relative increase of the difficulty of the task and shorter response times given that the verbal description was only presented once.

Methods

Participants

Thirty-four International Bachelor in Psychology students from a Dutch university participated for course credit ($M_{age} = 19.88$, $SD_{age} = 1.15$; 26 females).

Materials and apparatus

The same experimental task as in Experiment 1 was used. However, due to the low performance on trials that contained the cue "a person with a light colored skin", this formulation was changed into "a person with an olive colored skin".

Procedure

Participants received the same instructions as in Experiment 1, but with the additional instruction to not replay the verbal description after it had played once. The experimenter was in the same room as the participant and checked whether the participants followed this instruction. The session lasted approximately 20 min.

Data analysis

As in Experiment 1, 2 (visual complexity: low/high) x 2 (verbal ambiguity: low/high) repeated measures ANOVAs were conducted on the proportion correct and the average response times of correct trials. One participant failed to follow the instruction to listen to the verbal description only once and was excluded from all analyses. In addition, four participants were excluded from the response times (but not accuracy) analysis due to technical issues. This left 33 participants in the accuracy and 29 in the response times analysis.

Results and discussion

Data (Table 1) were analysed with 2 (visual complexity: low/high) x 2 (verbal ambiguity: low/high) repeated measures ANOVAs. On target identification accuracy, in contrast to our prediction this analysis did not reveal a main effect of visual complexity, $F(1, 32) < 1$, $p = 1.00$, $\eta^2_p = .00$. However, in line with our prediction there was a large main effect of verbal ambiguity, $F(1, 32) = 32.65$, $p < .001$, $\eta^2_p = .51$, indicating that participants had lower performance on high-ambiguity trials than on low-ambiguity trials. There was no interaction effect, $F(1, 32) < 1$, $p = .438$, $\eta^2_p = .02$.

On the response times, the same pattern of results was found. In contrast to our prediction there was no main effect of visual complexity, $F(1, 28) = 1.25$, $p = .273$, $\eta^2_p = .04$, but a large main effect of verbal ambiguity, $F(1, 28) = 26.27$, $p < .001$, $\eta^2_p = .48$, indicating that participants were slower to respond correctly on high-ambiguity trials than on low-ambiguity trials. There was no interaction effect, $F(1, 28) < 1$, $p = .983$, $\eta^2_p < .01$. In sum, the results of Experiment 2 revealed that participants were slower and less accurate in identifying the correct cartoon character for the high ambiguity trials than for the low ambiguity trials. However, the visual complexity of the images did not seem to influence performance. These results are at odds with the visual complexity results of Experiment 1 which showed that participants were slower in identifying the correct cartoon character for high visual complexity trials. One possible explanation for the difference in results regarding visual complexity might be due to the fact that participants in Experiment 2 were instructed not to replay the audio description, whereas participants of Experiment 1 were able to replay the audio description of the target. Other potential explanations for a lack of effect of visual complexity are discussed below.

General discussion

Research has shown that visual complexity and the ambiguity of verbal information affect the speed and accuracy of locating targets during visual search, but these factors are seldom studied in combination (even though they regularly co-occur, e.g. in multimedia learning materials). Therefore, the present study examined the potential interplay between the visual complexity of the stimuli and (verbal) instruction ambiguity on the performance of a visual search task (i.e. accuracy and speed). As predicted, higher verbal description ambiguity led to lower target identification accuracy and higher response times on accurate trials. This finding is in line with prior research in which verbal ambiguity resulted from non-specific location descriptions (Louwerse & Bangerter, 2010) or the visual contexts (Allopenna et al., 1998).

Contrary to our hypothesis, visual complexity did not affect accuracy or response times. The finding that visual complexity had little or no influence on target identification accuracy and response times deviates from earlier research and the results of Experiment 1. For instance, visual search became less efficient (Henderson, Chanceaux, & Smith, 2009; Wolfe, 1994) and less accurate (Neider & Zelinksky, 2011; Wolfe et al., 2002) when the target stimulus was embedded in a visually more complex background.

Possibly, the lack of effect of visual complexity on target identification in Experiment 2 results from the presence/nature of our verbal instruction (which was absent in the studies of Wolfe, 1994; Wolfe et al., 2002). Another possible explanation lies in the definition of visual complexity in our study (i.e. number of features contained within a display). High complexity trials contained more features and details than low complexity trials, which resulted in higher digital image file sizes, confirming a visual complexity difference (i.e. the more visually complex an uncompressed image, the more difficult it is to compress, the larger the image file; Chikhman et al., 2012; Donderi, 2006a; Marin & Leder, 2016). Nevertheless, a potential limitation of the present study is that this manipulation may have been too weak to yield complexity effects, although the subjective complexity ratings were higher for the high complexity images. Despite the results of the complexity ratings, some studies have shown discrepancies between objective features of visual complexity and perceived visual complexity (Yoon, Lim, & Ji, 2015). Similarly, in our study, we cannot rule out that the attributes that were added in the high-complexity displays (e.g. bottles, headphones) might have actually made the search process easier, whereas the relative homogeneity of characters in the low-complexity displays might have made it harder than intended. This explanation is also in line with the results of a study in which participants saw images of objects on which they performed both a categorisation task, requiring less differentiation of the objects, and an object decision task which required more differentiation of the objects (Gerlach & Marques, 2014). For the object decision task, visual complexity of the object images resulted in slower response times. Yet, for the object classification task, visual complexity of the object images resulted in faster response times, suggesting that visual complexity made the classification task easier rather than more difficult. In addition, Godwin et al. (2015) found that complex distractors were rejected faster than simple distractors, possibly due to the greater number of visual features available in the high complexity images to use to identify the target. Thus, although the high complexity stimuli were more complex as indicated by both an objective (file size) measure and subjective measure (complexity ratings), the lack of effects of the visual complexity manipulation in the current study might have been due to unintended effects of (dis)similarity of the cartoon images on the search process.

Regarding practical implications, our findings are relevant for the design of multimedia learning materials (e.g. instructional animations, videos). To facilitate learning from multimedia material, visual cues (e.g. pointing arrows or highlights) are often applied to guide the learner's attention towards the relevant information at the right time to enhance students' integration of the visual and verbal information (i.e. the signalling or cueing principle; Van Gog, 2014). Research findings regarding the effectiveness of these cues are mixed, however, with some finding beneficial effects of cueing (Jamet, 2014), whereas others do not (De Koning, Tabbers, Rikers, & Paas, 2011). Our findings suggest that verbal ambiguity may be a boundary condition for the effectiveness of cueing for learning. Although the present study used a task that is uncommon in multimedia learning environments, target identification based on verbal instructions is at the heart of many multimedia learning materials. The finding that high verbal ambiguity lowers target identification accuracy and increases response times suggests that visual cueing will be most effective for learning under high ambiguity conditions. Under high-ambiguity conditions, learners might not locate the referred information timely when there are no visual cues, which would hamper integration with the verbal description, and thereby, learning. High-ambiguity situations might arise when learners are confronted with new information and they lack the required prior knowledge to fully follow the multimedia learning materials even though visual cues are present. The latter ties in with the meta-analysis by Richter, Scheiter, and Eitel (2016), which showed that prior knowledge was a significant moderator of the cueing effect, with cueing being more helpful for low prior knowledge learners (for whom a verbal description would likely be more ambiguous). It would be interesting for future research to examine the potential role of verbal ambiguity in multimedia materials in combination with factors like the visual complexity of the material and the prior knowledge to gain more knowledge on how to make these learning materials more effective.

In conclusion, verbal description ambiguity, but not visual complexity, affected performance on a target identification task. These findings suggest that the ambiguity of verbal instructions can have a detrimental influence on how visual information is processed. These findings have implications, for instance, for the design of multimedia learning materials.

Acknowledgements

The authors would like to thank Susan Ravensbergen for assistance with data collection, and Jacqueline de Nooijer for recording the verbal descriptions.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was funded by a Vidi grant (# 452-11-006) from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organization for Scientific Research; NWO) awarded to Tamara van Gog.

ORCID

Tim van Marlen http://orcid.org/0000-0002-0894-989X

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*, 419–439. doi:10.1006/jmla.1997.2558

Castelhano, M. S., & Heaven, C. (2010). The relative ontribution of scene context and target features to visual search in scenes. *Attention, Perception, & Psychophysics*, *72*, 1283–1297. doi:10.3758/APP.72.5.1283

Chikhman, V., Bondarko, V., Danilova, M., Goluzina, A., & Shelepin, Y. (2012). Complexity of images: Experimental and computational estimates compared. *Perception*, *41*, 631–647. doi:10.1068/p6987

Davis, E. T., Shikano, T., Peterson, S. A., & Michel, R. K. (2003). Divided attention and visual search for simple versus complex features. *Vision Research*, *43*, 2213–2232. doi:10.1016/S0042-6989(03)00339-0

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2011). Attention cueing in an instructional animation: The role of presentation speed. *Computers in Human Behavior*, *27*, 41–45. doi:10.1016/j.chb.2010.05.010

Donderi, D. C. (2006a). An information theory analysis of visual complexity and dissimilarity. *Perception*, *35*, 823–835. doi:10.1068/p5249

Donderi, D. C. (2006b). Visual complexity: A review. *Psychological Bulletin*, *132*, 73–97. doi:10.1037/0033-2909.132.1.73

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, *24*, 409–436. doi:10.1007/BF02143160

Forsythe, A. (2009). Visual complexity: Is that all there is? *Engineering psychology and cognitive Ergonomics. LNCS*, *5639*, 158–166.

Gerlach, C., & Marques, J. F. (2014). Visual complexity exerts opposing effects on object categorization and identification. *Visual Cognition*, *22*, 751–769. doi:10.1080/13506285.2014.915908

Godwin, H. J., Walenchok, S. C., Houpt, J. W., Hout, M. C., & Goldinger, S. D. (2015). Faster than the speed of rejection: Object identification processes during visual search for multiple targets. *Journal of Experimental Psychology: Human Perception & Performance*, *41*, 1007–1020. doi:10.1037/xhp0000036

Henderson, J. M., Chanceaux, M., & Smith, T. J. (2009). The influence of clutter on real-world scene search: Evidence from search efficiency and eye movements. *Journal of Vision*, *9*, 1–8. doi:10.1167/9.1.32

Hout, M. C., & Goldinger, S. D. (2015). Target templates: The precision of mental representations affects attentional guidance and decision-making in visual search. *Attention, Perception, & Psychophysics*, *77*, 128–149. doi:10.3758/s13414-014-0764-6

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*, 151–171. doi:10.1016/j.actpsy.2010.11.003

Jamet, E. (2014). An eye-tracking study of cueing effects in multimedia learning. *Computers in Human Behavior*, *32*, 47–53. doi:10.1016/j.chb.2013.11.013

Louwerse, M. M., & Bangerter, A. (2010). Effects of ambiguous gestures and language on the time course of reference resolution. *Cognitive Science*, *34*, 1517–1529. doi:10.1111/j.1551-6709.2010.01135.x

Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision*, *9*, 1–13. doi:10.1167/9.11.8

Marin, M. M., & Leder, H. (2016). Effects of presentation duration on measures of complexity in affective environmental scenes and representational paintings. *Acta Psychologica*, *163*, 38–58. doi:10.1016/j.actpsy.2015.10.002

Neider, M. B., & Zelinksky, G. J. (2011). Cutting through the clutter: Searching for targets in evolving complex scenes. *Journal of Vision*, *11*, 1–16. doi:10.1167/11.14.7

Richter, J., Scheiter, K., & Eitel, A. (2016). Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis. *Educational Research Review*, *17*, 19–36. doi:10.1016/j.edurev.2015.12.003

Schmidt, J., & Zelinsky, G. J. (2009). Search guidance is proportional to the categorical specificity of a target cue. *Quarterly Journal of Experimental Psychology*, *62*, 1904–1914. doi:10.1080/17470210902853530

Van Gog, T. (2014). The signaling (or cueing) principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 263–278). New York, NY: Cambridge University Press.

Wolfe, J. M. (1994). Visual search in continuous, naturalistic stimuli. *Vision Research*, *34*, 1187–1195.

Wolfe, J. M., Oliva, A., Horowitz, T. S., Butcher, S. J., & Bompas, A. (2002). Segmentation of objects from backgrounds in visual search tasks. *Vision Research*, *42*, 2985–3004. doi:10.1016/S0042-6989(02)00388-7

Yang, H., & Zelinsky, G. J. (2009). Visual search is guided to categorically-defined targets. *Vision Research*, *49*, 2095–2103. doi:10.1016/j.visres.2009.05.017

Yoon, S. H., Lim, J. H., & Ji, Y. G. (2015). Perceived visual complexity and visual search performance of automotive instrument cluster: A quantitative measurement study. *Journal of Human–Computer Interaction*, *31*, 890–900. doi:10.1080/10447318.2015.1069661