Published in final edited form as:
Acta Psychol (Amst). 2008 October ; 129(2): 255–263. doi:10.1016/j.actpsy.2008.08.006.

Attention to Smoking-Related and Incongruous Objects during Scene Viewing

Verena S. Bonitz and **Robert D. Gordon**
North Dakota State University

Abstract

The present study examined influences of semantic characteristics of objects in real-world scenes on allocation of attention as reflected in eye movement measures. Stimuli consisted of full-color photographic scenes, and within each scene, the semantic salience of two target objects was manipulated while the objects' perceptual salience was kept constant. One of the target objects was either inconsistent or consistent with the scene category. In addition, the second target object was either smoking-related or neutral. Two groups of college students, namely current cigarette smokers (N = 18) and non-smokers (N = 19), viewed each scene for ten seconds while their eye movements were recorded. While both groups showed preferential allocation of attention to inconsistent objects, smokers also selectively attended to smoking-related objects. Theoretical implications of the results are discussed.

Keywords

visual attention; scene perception; cognitive bias; tobacco smoking

When we look at pictures of real-world scenes, we move our eyes from one location to another, gathering and processing the information that is presented to us. As a result of the large decrease in retinal acuity at higher visual eccentricities (Anstis, 1974), our eyes perform a series of rapid movements (saccades) in alternation with steady fixations in order to obtain visual information in optimal resolution (Carpenter, 1988). The resulting pattern of eye movements and fixations is not random, but mediated by the distinctiveness or informativeness of locations in the scene (e.g. Antes, 1974; Buswell, 1935; Mackworth & Morandi, 1967; Yarbus, 1967). Distinct regions may be those containing perceptually salient objects (e.g., Itti & Koch, 2000; Parkhurst, Law, & Niebur, 2002) or those containing objects that are semantically incongruous with the scene context (e.g., Loftus & Mackworth,

Address correspondence to: Verena S. Bonitz Department of Psychology Iowa State University Ames, IA 50011 U.S.A. Phone: (515) 294-8480 Fax: (515) 294-6424 vsbonitz@iastate.edu.

PsycINFO classification: 2323, 2346, 2990

NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

1978; Underwood & Foulsham, 2006) or that are relevant to the observer's goals (e.g., Hayhoe & Ballard, 2005; Henderson, Brockmole, Castelhano, & Mack, 2007).

A number of prominent theories of visual search and scene perception have argued that attention, and the eyes, are guided by an analysis of the perceptual salience of objects within the scene (e.g., Koch & Ullman, 1985; Treisman, 1988; Wolfe, 1994). Wolfe (1994), for example, proposed that preattentive scene processing, carried out in parallel across a scene, produces a set of "feature maps" which represent the locations of visual features (colors, orientations, etc) within the scene. Activations within each feature map reflect the perceptual salience of each feature, such that, for example, a distinctively-colored object will receive greater activation. Combining feature maps produces an overall activation map, which is used to guide attention to the most salient objects in a scene.

There has been a recent emergence of computational models of attention that formalize this approach and emphasize the importance of perceptual salience in guiding attention and the eyes (e.g., Itti & Koch, 2000; Parkhurst et al., 2002; Walther & Koch, 2006). Itti and Koch (2000), for example, describe a model in which a complex scene may be decomposed into feature maps, representing several features (e.g., color, intensity, edge orientation, etc) at multiple spatial scales. Combining feature maps yields a saliency map, a 2-dimensional representation of the perceptual salience of objects within the scene. Activations within the saliency map guide shifts of overt and covert attention. Attention is directed to the most salient object in the scene; that region of the saliency map is then inhibited, permitting serial deployment of attention from one region to the next most salient region in the scene. The model has been successful in predicting many aspects of human performance (e.g., Itti & Koch, 2000).

Nonetheless, considerable recent evidence suggests that overt and covert attention are not guided by perceptual salience alone when viewing complex, meaningful scenes (e.g., Becker, Pashler, & Lubin, 2007; Henderson et al., 2007; Land & Hayhoe, 2001; Navalpakkam & Itti, 2005; Torralba, Oliva, Castelhano, & Henderson, 2006; Underwood & Foulsham, 2006; Underwood, Humphreys, & Cross, 2007). Instead, a number of cognitive factors have been shown to influence shifts of attention during scene viewing. Land and Hayhoe (2001), for example, have shown that eye movements are guided by task constraints rather than by perceptual salience, such that task-relevant objects are more likely to be fixated than less relevant objects. Furthermore, a number of studies have suggested that shifts of attention during visual search are constrained by knowledge of likely target locations within the scene (e.g., Brockmole, Castelhano, & Henderson, 2006; Chun & Jiang, 1998). Finally, some evidence suggests that attention is guided by an analysis of the semantic properties of the scene and of the objects in contains, although this conclusion has been controversial (De Graef, 2005; De Graef, Christiaens, & d'Ydewalle, 1990; Gordon, 2004, 2006; Henderson, Weeks, & Hollingworth, 1999; Loftus & Mackworth, 1978; Underwood & Foulsham, 2006; Underwood et al., 2007).

An influential early study in which semantic informativeness was directly manipulated was conducted by Loftus & Mackworth (1978). The stimuli consisted of line drawings that contained a target object which was either likely to appear within the scene context (consistent object) or unlikely to appear (inconsistent object). For example, a tractor is likely to be found in a farm scene, whereas an octopus is not. Likewise, the octopus, but not the tractor, is consistent with an underwater scene. Each scene was presented for four seconds, and participants were instructed to memorize the scenes for a later memory test. A main result of this study was the tendency of participants to fixate inconsistent objects earlier than consistent objects following the first fixation on the scene. This result, however, has not been replicated in later studies (De Graef et al., 1990; Henderson et al., 1999). Henderson et al. (1999), for example, conducted a systematic study of semantic influences on eight measures of target processing time reflecting the allocation of overt attention via eye movements. These included the first and second-pass gaze durations, first and second-pass fixation count, total fixation duration, and total fixation count. Participants viewed line drawings containing either a consistent or an inconsistent object for 15 seconds (in preparation for a later memory test) while their eye movements were monitored. Henderson et al. reported a significant influence of object consistency on processing time. Inconsistent objects were fixated longer and more frequently than consistent objects, and participants' eye movements returned to these informative regions more often. However, unlike Loftus and Mackworth (1978), Henderson et al did not find that inconsistent objects were fixated earlier than consistent objects.

The inconsistency across studies has been explained as a result of differences in the complexity and size of scenes (Henderson & Hollingworth, 1998). Another explanation is based on visual crowding effects (Becker et al., 2007); while most studies have used line-drawn scenes, studies in which scenes were much more densely populated with objects have not reported earlier fixation of inconsistent objects (e.g., Henderson et al., 1999). When scenes are densely populated, the peripheral target objects of high spatial frequency may be especially vulnerable to masking, reducing the likelihood of those objects drawing attention in the later studies.

A significant drawback in the reviewed literature on object consistency in scene perception is that changes to an object's visual and semantic characteristics have, in many cases, been confounded. For example, when a tractor is replaced by an octopus in a farm scene (Loftus & Mackworth, 1978), not only is the semantic context changed, but also the visual attributes of the target object and their integration with the surrounding features; a tractor has a rather rectangular shape, whereas an octopus is relatively round. Similarly, an example of a scene used by Gordon (2004), contains a live chicken in a kitchen scene as an inconsistent object. The chicken's round and smooth features contrast with the grid-like pattern of horizontal and vertical lines of the kitchen furnishing, and attention might have been drawn towards the object on the basis of visual informativeness.

Therefore, one goal of the present study was to investigate whether semantic properties of objects influence attentional allocation in scenes when size and visual properties of semantically consistent and inconsistent objects were matched. Previous research suggests that this may be the case. Underwood et al. (2007), for example, equated the perceptual salience of consistent and inconsistent objects, and found that inconsistent objects were fixated earlier than consistent objects. Participants in that study inspected scenes in preparation for a memory test. Given that the distribution of attention during scene viewing depends on the viewer's task (Land & Hayhoe, 2001; Underwood, Foulsham, van Loon, Humphreys, & Bloyce, 2006) and that preparation for a memory test may encourage participants to attend to the most unique scene elements, we sought to extend those findings to a situation in which participants freely view scenes without anticipating a memory test. In addition, full color photographs were used as stimuli, in contrast to the black and white line drawings used in many previous studies (Gordon, 2004; Henderson et al., 1999; Loftus & Mackworth, 1978) in order to make the stimuli appear more realistic, and to reduce visual crowding effects. In black and white line drawings, object recognition depends on high frequency contour information. Therefore, objects in the periphery might not be identifiable, especially in highly cluttered scenes (Becker et al., 2007). In full color photographs, however, additional routes to object identification (color contrast, diagnostic colors of objects, etc) are available (Oliva & Schyns, 2000), and visual crowding should therefore be less of a concern.

Individual Differences in Attentional Bias

Much prior research has been focused on a general bias to attend to objects based on their identities. It is likely, however, that attention during scene viewing may also be guided by individual preferences and biases toward personally relevant objects. It has been observed, for example, that subjects from certain clinical populations exhibit an attentional bias towards stimuli that relate to their pathologies. Attentional biases have been demonstrated to be part of the cognitive structure of various disorders, including eating disorders (e.g., Lee & Shafran, 2004), anxiety and depression (e.g., Mogg & Bradley, 2005), and addictive behaviors such as cigarette smoking (Bradley, Mogg, Wright, & Field, 2003; Field, Mogg, & Bradley, 2004; Mogg & Bradley, 2002; Mogg, Bradley, Field, & De Houwer, 2003; Waters & Feyerabend, 2000).

Attentional bias for smoking-related stimuli is an important factor underlying addiction maintenance: For example, attentional bias has been shown to be associated with self-reported urge to smoke (Mogg & Bradley, 2002; Mogg et al., 2003), craving (Mogg et al., 2003), nicotine deprivation (Field et al., 2004), number of previous quit attempts (Bradley et al., 2003), and frequency of smoking (Mogg, Field, & Bradley, 2005). In addition, greater strength of attentional bias has been found to be associated with a higher risk of relapse after attempting to quit (Waters et al., 2003).

Many experiments on attentional bias have been conducted in order to learn more about the cognitive mechanisms and different attentional components involved in addictive behaviors such as cigarette smoking. One important theoretical aspect discussed in the literature is the distinction between initial orienting towards smoking cues and maintenance of attention once the stimulus has been attended to (e.g., Field, Mogg, & Bradley, 2006; LaBerge, 1995). Visual probe tasks have been used to demonstrate this distinction by manipulating Stimulus Onset Asynchrony (SOA).

Whereas the evidence seems to be consistent across studies with regard to attention maintenance, the results are mixed concerning the initial orientation bias towards smoking cues (e.g., Bradley et al., 2003; Field et al., 2006; Mogg et al., 2003). For example, a visual probe task was conducted with 20 paired sets of photographs (Bradley et al., 2003). One set of pictures showed a smoking related scene (e.g., a man holding a cigarette); the other set was closely matched for content, but did not contain any smoking related cues (e.g., a man holding a pen). A pair of pictures was shown for 500 ms (assessing initial orientation) and 2,000 ms (assessing attention maintenance), respectively, and a probe that participants had to identify as quickly as possible appeared at the location of one of the two pictures shortly after stimulus offset. In the 2,000 ms condition, smokers responded faster to the probe replacing the smoking-related picture, indicating attentional bias. In the 500 ms condition, in contrast, only the subgroup of smokers who had a history of a high number of unsuccessful quit attempts showed vigilance towards the smoking-related stimuli. Non-smoking control participants reacted equally fast to smoking-related and neutral pictures.

In a next step, participants performed the visual probe task while their eye movements were recorded (Mogg et al., 2003). Probability of initial fixation and overall gaze duration were calculated for both smoking-related and neutral pictures. The results revealed that smokers were more likely to place their initial fixation on the smoking-related pictures, and gaze durations were significantly longer for these pictures. For the non-smoking control group, there was no difference in eye movement measures for smoking-related and neutral pictures.

One limitation of Mogg et al.'s (2003) study was that only the distribution of attentional resources between two photographs (where one of them contained a smoking-related cue) was assessed, but not the allocation of attention to different objects within one single scene, making it impossible to study the interaction of an object with the scene gist. In addition, the photographs used as stimuli were rather simple in content; most often, only a single object was depicted in front of a neutral background (M. Field, personal communication, November 16, 2005).

Therefore, the second goal of the present study was to address this limitation, and to test whether the attentional bias effect (and the differentiation between initial orienting and maintenance) could be replicated with stimuli that were more complex and realistic. Although natural scenes comparable to those in the present study have been used in a flicker paradigm to assess attentional bias in social users of alcohol and cannabis (Jones, Bruce, Livingstone, & Reed, 2006; Jones, Jones, Smith, & Copley, 2003), this type of stimulus has not been used in an eye movement study to assess attentional bias in cigarette smokers.

The Present Study

In the present study, two independent subject groups (smokers and a non-smoking control group) were tested. Participants viewed scenes in which two target objects had been manipulated. First, each scene contained an object that was either consistent or inconsistent with the scene context. Second, the scene also included either a smoking-related object (e.g., a cigarette, cigarette packet, lighter, or an ashtray filled with cigarette butts), or a neutral object with similar visual features. Importantly, both smoking-related and neutral object types were consistent with the scene context. Hence, four object combinations were used in each scene category: consistent/neutral, inconsistent/neutral, consistent/smoking-related, inconsistent/smoking-related (see Figure 1).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

The main hypotheses of the study are the following: In line with the published findings on object consistency in scenes, it was expected that semantically inconsistent target objects would be fixated longer and more often than consistent objects (Henderson et al., 1999); this inconsistent object advantage should occur for both smokers and non-smokers alike. If, as Loftus & Mackworth (1978) argue, attention is drawn to the location of inconsistent objects, then both groups would be expected to fixate inconsistent objects earlier than consistent objects. If, on the other hand, Henderson et al. (1999) are correct that attention is guided primarily by perceptual object properties, then no difference would be predicted in the latency with which inconsistent and consistent objects are fixated.

Smokers, as predicted by Mogg et al.'s (2003) results, were expected to fixate smoking-related objects longer, and to refixate such objects more often, than a neutral target object with similar visual features. No prediction could be made for number of prior fixations before target fixation to smoking-related objects because the empirical evidence to date is inconclusive. No differences in eye movement patterns between smoking-related and neutral target objects were expected for non-smokers.

Method

Participants

Eighteen current smokers (six male) and 19 current non-smokers (four male) recruited at North Dakota State University participated in the study for either course credit or a monetary compensation of $ 10/hour. Two of the non-smokers had quit smoking ten and 16 years prior to the date of the experiment, respectively. Evidence from an ERP study indicates that ex-smokers react towards smoking cues in a manner similar to non-smokers after a period of six months of abstinence (Littel & Franken, 2007). For the group of smokers, the average number of cigarettes smoked per day was 10.7 (SD = 5.2), and at the time of the experiment they had been smoking for an average of 8.2 years (SD = 6.3). All participants had normal or corrected-to-normal vision.

Stimuli and Apparatus

Full-color photographs (subtending 32° by 45°) of real-world scenes were used as stimuli. There were ten different scenes, and each scene was staged and photographed in four versions, accounting for four object types (see Figure 1). The first manipulation concerned the semantic consistency of the object; a scene could either contain a scene-consistent or a scene-inconsistent object, both objects matched for size and visual features (mean visual angle of consistent objects: 7.0° (vertical) and 9.0° (horizontal); mean visual angle of inconsistent objects: 6.7° (vertical) and 8.8° (horizontal)). As the second manipulation, one scene version contained a smoking-related object (mean visual angle 4.9° (vertical) and 5.5° (horizontal)), which was replaced by a neutral (i.e., smoking-unrelated) object (mean visual angle: 5.1° (vertical) and 5.6° (horizontal)). Both neutral and smoking-related object types were semantically consistent with the scene context. A list of all ten scene categories and objects manipulated is presented in the Appendix.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

In order to demonstrate that each pair of objects had similar visual saliency, a visual saliency map was constructed for each scene using the Saliency Toolbox for Matlab (Walter & Koch, 2006). Matlab was used to determine intensities across the saliency map (ranging from 0 to 1, with higher intensities indicating higher salience). The result of this analysis indicated that visual saliency was similar for the respective object pairs. Indeed, to the extent that there was a difference between consistent and consistent objects, *consistent* objects tended to be slightly more visually salient ($M = 0.470$) than inconsistent objects ($M = 0.391$), though this difference was not significant, $t(19) = 1.309$, $p > 0.05$. There was no difference in salience between smoking related objects ($M = 0.349$) and neutral objects ($M = 0.332$), $t(19) = 0.421$, $p > 0.05$.

In addition to the experimental scenes, six filler scenes were included. In these filler scenes two neutral objects were replaced by visually matching objects following the same procedure as in the experimental scenes. The filler scenes did not contain any inconsistent or smoking-related objects, based on the rationale that, if every scene contained such object types, participants might start explicitly searching for them.

The photographs were presented full screen on a NEC MultiSync FP2141SB monitor at a resolution of 1024 × 768 pixels at a refresh rate of 75 Hz, and at a viewing distance of 57 cm. A head-mounted EyeLink II eyetracker (SR Research Ltd., Mississauga, Ontario, Canada) was used to record the participants' eye movements, sampling pupil position at a rate of 500 Hz.

Procedure

Upon their arrival at the laboratory, participants provided informed consent, followed by the completion of a questionnaire on smoking habits (NIH, 2006) and attitudes (Shore, Tashchian, & Adams, 2000). Participants were then fitted with the eye-tracking equipment, and began the experimental trials. At the beginning of each trial a fixation cross (1° visual angle) appeared in the center of the screen. By pressing a key participants initiated a drift correction, which was followed by the full-screen appearance of one of 64 scenes. The scene order was randomized for each participant. The scene was visible for 10 seconds during which participants were instructed to freely examine the scene while their eye movements were recorded. After scene offset, participants rated the pleasantness of the scene they just saw on a scale from 1 (unpleasant) to 7 (pleasant) by pressing a key, which was a procedure adapted from Mogg et al. (2003). The inclusion of this task was intended to orient participants to the scene during the viewing period. The response then initiated the appearance of the fixation cross for the next trial. Participants completed 64 trials; the experimental session lasted approximately 40-50 minutes.

Results and Discussion

The following eye movement parameters, as a measure of the allocation of overt attention, were the dependent variables in this study: (a) number of total discrete fixations on the target object (total fixation count); (b) total duration of all fixations on the target object (dwell time); (c) duration of the first fixation on the target object (first fixation duration); (d) probability of target object fixation (proportion of the ten experimental scenes in which the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

target object was fixated at least once – for example, if the consistent object was fixated in nine out of the ten scenes, the fixation probability was .90); (e) number of fixations prior to fixating the target object.

The eye movement data were analyzed with EyeLink Data Viewer software (SR Research Ltd., Mississauga, Canada). The Data Viewer software identifies blinks and saccades (defined as eye velocities exceeding 30 deg/s or eye acceleration exceeding 8000 deg/s^2); periods that are not classified as blinks or saccades are considered discrete fixations in the analyses below. For each target object pair an interest area was defined by using the outer target object contours as boundaries. A fixation was allocated to a given object when it occurred in a spatial location within the interest area circumscribed by the object contours.

The results from the ten experimental scenes were averaged for further analysis. If a target object had not been fixated at all in one of the ten scenes, the data pair for the respective object type combination for this particular scene was excluded from the analysis for the particular subject, and the results were averaged across the remaining scenes. This procedure was only applied in a minority of the trials, since all target objects were fixated in at least eight out of ten scenes or more (see Table 1 for the average fixation probabilities for each target object). A second issue was that, for one scene, the location of a target object coincided with the location of the initial fixation cross; hence, we did not analyze the number of fixations prior to fixating the target object for the scene in question. As a result of these issues, a total of 7.7% of data values were omitted.

The experimental design included three independent variables with two levels each. “Group” was the between-subjects factor, the two levels consisting of smokers and non-smokers. “Target object” was a within-subjects factor. In one set of analyses, the target object contrast was between inconsistent and consistent target objects, while in the other set of analyses the contrast was between neutral and smoking-related target objects. Each of the target objects could also occur in two different contexts: Both consistent and inconsistent objects could appear in the presence of either a neutral or smoking-related object, and vice versa. Therefore, “context” was the second within-subjects factor. The data were analyzed by two 2 (group) x 2 (target) x 2 (context) repeated-measures MANOVAs with five dependent variables, each followed up by univariate ANOVAs. The first analysis looked at the semantic consistency contrast, the second at the contrast between smoking-related and neutral objects. Mean values and standard deviations for each dependent variable are displayed in Table 1.

Object consistency manipulation

The main hypothesis for the object consistency manipulation was that participants would show preferential allocation of attention to semantically inconsistent objects; that is, a significant main effect of target object was expected. This hypothesis was confirmed (Wilk’s Lambda = .212, $F(5,31) = 23.0$, $p < .001$). More specifically, univariate follow-up analyses revealed that a semantic consistency effect could be observed in all five dependent variables: Inconsistent objects were fixated significantly more often ($F(1,35) = 52.7$, $p < .001$) and longer ($F(1,35) = 43.7$, $p < .001$) than consistent objects. In addition, the first fixation duration was significantly longer for inconsistent objects ($F(1,35) = 5.9$, $p = .02$), and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

participants had significantly fewer fixations prior to target object fixation for the inconsistent objects ($F(1,35) = 28.7$, $p < .001$). Moreover, participants were significantly more likely to fixate inconsistent objects relative to consistent objects ($F(1,35) = 28.9$, $p < .001$).

As expected, there was no significant main effect of group (Wilk's Lambda = .923, $F(5,31) = 0.52$, $p = .76$), as smokers and non-smokers were similar in their patterns of attention allocation to semantically consistent and inconsistent objects. In addition, all interaction terms remained below statistical significance (Wilk's Lambdas > .90, $Fs(5,31) < 1$, $ps > .05$).

The manipulation of object consistency in the presence of a neutral object in our study was similar in many ways to the designs reported previously in the scene perception literature (e.g., De Graef et al., 1990; Henderson et al., 1999; Loftus & Mackworth, 1978; Underwood et al., 2007), and the results can be compared accordingly. Our finding of higher fixation counts, longer dwell times and first fixation durations for inconsistent objects replicate the effects presented in these studies. This indicates that the processing of semantically inconsistent objects (e.g., recognition and integration with the overall scene category) is more effortful and time demanding relative to objects consistent with the scene category. In addition, we found that inconsistent objects were fixated earlier than consistent objects. This result is consistent with Loftus and Mackworth's (1978) study and with some more recent work (e.g., Underwood et al., 2007), but inconsistent with the other findings (De Graef et al., 1990; Gareze & Findlay, 2007; Henderson et al., 1999). Possible explanations to reconcile these findings will be described in the General Discussion.

Smoking cue manipulation

When considering the overall MANOVA effect for the contrast between smoking-related and neutral objects, the result was a significant main effect of target object (Wilk's Lambda = .642, $F(5, 31) = 3.45$, $p = .014$), and a significant main effect of context (Wilk's Lambda = .618, $F(5, 31) = 3.83$, $p = .008$). Neither the main effect of group (Wilk's Lambda = .731, $F(5, 31) = 2.28$, $p = .07$), nor the group by target interaction (Wilk's Lambda = .790, $F(5, 31) = 1.65$, $p = .18$) reached statistical significance.

Univariate follow-up tests were conducted to obtain a better understanding of the effects. There was a significant main effect for all dependent variables ($Fs(1, 35) > 4.17$, $ps < .05$) except for the number of prior fixations before target object fixation ($F(1, 35) = 0.84$, $p = .37$): A comparison of the means revealed that participants overall were more likely to fixate the smoking-related objects than the neutral objects, the smoking-related objects were fixated longer and more often, and received longer first fixation durations. This main effect, however, was qualified by a significant target object by group interaction for the fixation count ($F(1, 35) = 4.06$, $p = .05$) and dwell time ($F(1, 35) = 6.23$, $p < .05$) variables. Based on the means, smokers spent on average 1,644 ms (SD = 87 ms) fixating smoking-related objects, but neutral objects were fixated only for an average of 1,258 ms (SD = 72 ms). For the group of non-smokers, dwell times for smoking-related objects (M = 1,451 ms, SD = 85 ms) did not differ significantly from dwell times for neutral objects (M = 1,417 ms, SD = 70 ms). The results were similar for the fixation count variable: smokers fixated the smoking-related objects on average 5.1 times (*SD* = 0.3) during the ten seconds the scenes were presented, and the neutral objects received an average of 4.1 fixations (*SD* = 0.3). Again, for non-smokers, smoking-related objects (*M* = 4.8, *SD* = 0.3) were fixated equally often as neutral objects (*M* = 4.7, *SD* = 0.3). These results show that smokers, but not non-smokers, showed an attentional bias toward smoking cues. More specifically, the selective bias pertained to the maintenance of attention (as operationalized in terms of overall fixation count and dwell time measures). As mentioned above, there was no indication that either group showed initial orienting bias towards smoking-cues (the number of fixations prior to target fixation for both groups was similar for smoking-related and neutral objects). These findings from the present study are consistent with previous studies (e.g., Bradley et al., 2003; Field et al., 2004; Field et al., 2006; Mogg et al., 2003).

Another interesting finding was the significant main effect of context found in the overall MANOVA. Univariate follow-up analyses revealed that this overall effect was due to significant differences in dwell time ($F(1, 35) = 5.87$, $p = .02$), fixation count ($F(1, 35) = 7.35$, $p = .01$), and number of fixations prior to target fixation ($F(1, 35) = 4.12$, $p = .05$), depending on whether the target objects appeared concurrently with a consistent or inconsistent object. A comparison of the means (Table 1) shows that smoking-related and neutral objects were fixated longer and more often when they appeared in the presence of a consistent object compared to when appearing in the presence of an inconsistent object. Moreover, the target objects were fixated earlier in scene viewing when a consistent object was present. Taken together, these findings suggest that semantically inconsistent objects were more salient for participants than either the neutral or smoking-related objects, hereby drawing attentional resources at the expense of these objects. Further support for this mechanism comes from the finding that the effects were much larger for the semantic consistency manipulation than for the smoking cue manipulation.

General Discussion

Object consistency

We found that participants exerted significantly fewer fixations prior to fixating inconsistent objects in comparison to consistent objects. This result is consistent with some previous studies (e.g., Loftus & Mackworth, 1978; Underwood et al., 2007), but inconsistent with others (e.g., De Graef et al., 1990; Henderson et al., 1999). Two main factors might contribute to this discrepancy. The first concerns the nature of the stimuli. The scenes used in many previous studies were black and white line drawings, whereas the scenes used in the present study were full color photographs. One reason why Henderson et al. (1999) and De Graef et al. (1990) might have failed to find any differences is the effect of visual crowding, as pointed out by Becker et al. (2007). A cluttered image can make it difficult to identify target objects, especially those in the periphery, when viewing line drawings, in which object recognition is based on contour information alone. Therefore, object identification might have been restricted to a small region surrounding a fixated scene location. Importantly, the fact that an effect of object consistency on the number of fixations prior to first fixation was found in Loftus and Mackworth's study, whose scenes contained comparatively few objects but a lot of free space (Henderson et al., 1999), supports the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

hypothesis that the degree of visual crowding might play a role in whether or not inconsistent objects are fixated earlier than consistent objects or not. It is worth noting, however, that a more recent study using photographic scenes (Gareze & Findlay, 2007) failed to find evidence that inconsistent objects are fixated earlier than consistent objects.

Another factor that might explain the inconsistency across studies with regard to the number of fixations prior to target fixation is the insufficient differentiation of visual and semantic saliency in many earlier studies. For example, in Henderson et al.'s (1999) study, two objects were swapped across scene categories, thereby creating the inconsistent object conditions. A limitation of this procedure is that an object's visual salience is partially defined by its surroundings; if, for example, a round object is introduced into a grid-like location containing primarily straight lines, the object will visually stand out more than if it is inserted among similarly round-shaped objects. As a result, a semantically consistent object might have been more visually salient than an inconsistent object that replaced it in the same scene (Becker et al., 2007). These uncontrolled effects of visual salience could have masked any effects of object consistency. The present study, however, was designed explicitly to control for low-level visual factors, which might otherwise attract attention, by matching the target object pairs in terms of their visual characteristics. Therefore, the observed inconsistent object advantage should be due to the object's semantic characteristics alone. Equating perceptual salience of consistent and inconsistent objects may be critical to observing this inconsistent object advantage; indeed, the most recent study to report such an advantage for inconsistent objects (Underwood et al., 2007) also used target objects that had been matched for perceptual salience.

Two competing accounts have been described that offer an explanation for why individuals direct more attentional resources towards semantically inconsistent objects (see Gordon, 2006 for a detailed description). According to the semantic conflict interpretation (e.g., Gordon, 2004; Loftus & Mackworth, 1978), attention is drawn to inconsistent objects because these objects are in semantic conflict with the overall scene gist, and cognitive resources are needed to resolve this conflict. This view assumes that participants easily identify the inconsistent object, but have difficulty integrating the object with the representation of the scene in which it is embedded. On the other hand, the local processing difficulty interpretation (e.g., Johnston, Hawley, Plewe, Elliott, & DeWitt, 1990) asserts that attention is drawn to objects that are difficult to identify (and that therefore represent "trouble spots" needing additional resources). In the case of semantically inconsistent objects, individuals do not have the contextual information needed to rapidly identify these objects. Prior research has shown that the overall scene gist can have an effect on object perception and identification (Biederman, Mezzanotte, & Rabinowitz, 1982; Boyce & Pollatsek, 1992; Davenport & Potter, 2004; Gordon, 2004, 2006). More specifically, identification of objects is facilitated when the target object is semantically consistent with the overall scene gist. Inconsistent objects, which by definition do not benefit from such facilitation, are therefore more likely to represent areas of local processing difficulty, and to become the focus of attention. The local processing difficulty hypothesis has been supported experimentally by Gordon (2006) in a series of experiments assessing the priming effect of semantically consistent and inconsistent objects in a lexical decision task.

In the present study object identification was not assessed, and it is therefore difficult to use the present results to evaluate these competing hypotheses. For example, the finding of fewer number of fixations prior to target object fixation, longer first fixation durations, and higher fixation probabilities for inconsistent objects could implicate both object identification as well as semantic integration processes. However, the fact that participants returned their gaze to the inconsistent objects more often than to the consistent objects might be interpreted as efforts towards resolution of the semantic conflict rather than object identification difficulties; if object identification difficulty was the only mechanism underlying the semantic consistency effect, participants would not have a large incentive to return to a previously attended region, since object identification has likely been completed. Thus, this result may suggest difficulties in integrating inconsistent object representations later in scene viewing, though the initial orientation of attention to the inconsistent objects may reflect difficulties in either integrating or identifying those objects. More research is needed to obtain a better understanding of the mechanism(s) underlying the object inconsistency effect in the present experimental paradigm.

Attentional bias towards smoking cues

The comparison of the results for neutral and smoking-related objects supported the existence of an attentional bias of smokers towards smoking cues. More specifically, this effect was found in the two measures of attention maintenance (fixation count and dwell time). This finding is consistent with comparable previous studies (Bradley et al., 2003; Field et al., 2004; Field et al., 2006; Mogg et al., 2003). In contrast, the present study did not support the hypothesis of an initial orientating bias towards smoking cues. While it seems that, based on previous findings, attention maintenance is a robust effect that holds true across different experimental conditions, the initial orienting bias might be more susceptible to moderation by other variables. For example, Bradley et al. (2003) found that only those smokers who have had a large number of previous quit attempts showed the initial orienting bias, but not the group of smokers as a whole. In addition, strength of attentional bias in general has been shown to be associated with other variables which might function as possible moderators, such as self-reported urge to smoke (Mogg & Bradley, 2002; Mogg et al., 2003), craving (Mogg et al., 2003), duration of abstinence (Field et al., 2004), and mood (Bradley, Garner, Hudson, & Mogg, 2007). More research is needed that focuses on the identification of moderators of smokers’ attentional bias towards smoking cues, in particular with regard to the question of biases in initial orientation.

The finding of a significant main effect of target object concerning the first fixation duration and fixation probability variables (participants irrespective of group tended to be more likely to fixate smoking-related objects relative to neutral objects, and the average first fixation duration was longer for the smoking-related objects) might be an artifact of the experimental procedure. Participants completed the smoking questionnaire prior to the experimental task, which might have primed them to selectively attend to the smoking-related objects. In addition, although only about one third (20 out of 64) of the scenes contained any smoking-related objects, participants might have been led to selectively search for smoking-related objects in the scenes. That these effects are possible has been shown previously in a change detection task in which awareness of the focus on smoking of the experiment had been manipulated (Yaxley & Zwaan, 2005). These issues are a clear limitation of the current study. Nonetheless, the finding that non-smokers (in contrast to smokers) did not show maintenance bias towards the smoking cues despite possible priming and selective search for smoking-related objects can be taken as an indicator of the robustness of the maintenance effect.

Semantic consistency bias vs. attentional bias towards smoking cues

A unique feature of the study was the concurrent manipulation of semantically consistent or inconsistent objects and smoking-related or neutral objects within the same scene. This design allowed studying the relative saliency of the different object types. Two findings emerged: the first point concerns the distinction between initial orienting of attention and attention maintenance; the second point concerns the size of the respective effects. Whereas participants showed both initial orientation of attention and attention maintenance favoring the semantically inconsistent objects, attentional bias of smokers towards smoking cues was limited to attention maintenance only. This suggests that biases to smoking-related objects on the one hand and to semantically inconsistent objects on the other hand reflect different mechanisms underlying the allocation of attention. Second, the semantic inconsistency effect seems to be larger than the individual bias effect. Moreover, participants paid less overall attention to both neutral and smoking-related objects when they appeared in the presence of an inconsistent object as compared to when a consistent object was present.

Summary

In the present study we have shown that viewers of real-world scenes preferentially allocate attention to objects that are semantically salient. Both smokers and non-smokers attend earlier and more often to objects that are inconsistent with the scene context, but smokers also preferentially attend to smoking-related objects. The results from the present study confirm the existence of an attentional bias for smoking cues in smokers during the viewing of real-world scenes. Interestingly, this effect was limited to maintenance of attention once the object was fixated; smoking-related objects were not fixated earlier than the neutral objects. In summary, the current research has demonstrated that when we view complex natural scenes, we are biased towards attending to some objects rather than others. These biases can be general, as in the case of objects that are inconsistent with the overall scene category, but they can also reflect individual differences, as in the case of cigarette smokers who preferentially attend to smoking-related cues.

Acknowledgments

This project was supported by Grant Number 1P20 RR020151 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NCRR or NIH. The project was also supported by the National Science Foundation under Grant Number 01322899 to the second author. We thank Peter De Graef, Matt Field, Johan Wagemans, and an anonymous reviewer for helpful comments on the manuscript.

Appendix: Target Objects Manipulated in the Experimental Scenes

Scene	Consistent	Inconsistent	Smoking-related	Neutral
Billiard table	Billiard ball	Tomato	Cigarette pack	Cell phone
Dining table	Fork	Wrench	Lighter	Chewing gum
DVD setup	DVD	Make-up kit	Ashtray w/ stubs	Cup w/ snacks
Chemistry Lab	Erlenmeyer	Mike's liquor	No-smoking sign	No-food sign
Presenter stand	Key chain	Ice-cream	Cigarette	Pencil
Window sill	Sea shell	Make-up pad	Cigarette stub	Chewed gum
Study table	Apple	Softball	Cigarette pack	Deck of cards
Office desk	Paper	Mushroom	Cigarette pack	Camera
Bathroom sink	Toilet bag	Toaster	Cigarette pack	Cotton swabs
Tool shop	Screwdriver	Pickle	Lighter	Adhesive flask

References

Anstis SM. A chart demonstrating variations in acuity with retinal position. Vision Research. 1974; 14:589–592. [PubMed: 4419807]

Antes JR. The time course of picture viewing. Journal of Experimental Psychology. 1974; 103:62–70. [PubMed: 4424680]

Becker MW, Pashler H, Lubin J. Object-intrinsic oddities draw early saccades. Journal of Experimental Psychology: Human Perception and Performance. 2007; 33(1):20–30. [PubMed: 17311476]

Biederman I, Mezzanotte RJ, Rabinowitz JC. Scene perception: detecting and judging objects undergoing relational violations. Cognitive Psychology. 1982; 14:143–177. [PubMed: 7083801]

Boyce SJ, Pollatsek A. Identification of objects in scenes: The role of scene background in object naming. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1992; 18:531–543.

Bradley BP, Garner M, Hudson L, Mogg K. Influence of negative affect on selective attention to smoking-related cues and urge to smoke in cigarette smokers. Behavioural Pharmacology. 2007; 18(4):255–263. [PubMed: 17551317]

Bradley BP, Mogg K, Wright T, Field M. Attentional bias in drug dependence: Vigilance for cigarette-related cues in smokers. Psychology of Addictive Behaviors. 2003; 17(1):66–72. [PubMed: 12665083]

Brockmole JR, Castelhano MS, Henderson JM. Contextual cueing in naturalistic scenes: Global and local contexts. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2006; 32:699–706.

Buswell, GT. How people look at pictures. Univ. Chicago Press; Chicago: 1935.

Carpenter, RHS. Movements of the eyes. Pion; London: 1988.

Chun MM, Jiang Y. Contextual cueing: Implicit learning and memory of visual context guides spatial attention. Cognitive Psychology. 1998; 36:28–71. [PubMed: 9679076]

Davenport JL, Potter MC. Scene consistency in object and background perception. Psychological Science. 2004; 15(8):559–564. [PubMed: 15271002]

De Graef, P. Semantic effects on object selection in real-world scene perception. In: Underwood, G., editor. Cognitive processes in eye guidance. Oxford University Press; Oxford: 2005. p. 189-212.

De Graef P, Christiaens D, d'Ydewalle G. Perceptual effects of scene context on object identification. Psychological Research. 1990; 52:317–329. [PubMed: 2287695]

Field M, Mogg K, Bradley BP. Eye movements to smoking-related cues: Effects of nicotine deprivation. Psychopharmacology. 2004; 173(1-2):116–123. [PubMed: 14663552]

Field, M.; Mogg, K.; Bradley, BP. Attention to drug-related cues in drug abuse and addiction: Component processes. In: Wiers, RW., editor. Handbook of implicit cognition and addiction. Sage Publications, Inc.; Thousand Oaks, CA: 2006. p. 151-163.

Gareze, L.; Findlay, JM. Absence of scene context effects in object detection and eye gaze capture. In: van Gompel, RPG., editor. Eye movements: A window on mind and brain. Elsevier; Amsterdam: 2007. p. 617-637.

Gordon RD. Attentional allocation during the perception of scenes. Journal of Experimental Psychology: Human Perception & Performance. 2004; 30(4):760–777. [PubMed: 15301623]

Gordon RD. Selective attention during scene perception: Evidence from negative priming. Memory & Cognition. 2006; 34(7):1484–1494. [PubMed: 17263073]

Hayhoe M, Ballard D. Eye movements in natural behavior. Trends in Cognitive Sciences. 2005; 9:188–193. [PubMed: 15808501]

Henderson, JM.; Brockmole, JR.; Castelhano, MS.; Mack, M. Visual saliency does not account for eye movements during visual search in real-world scenes. In: van Gompel, RPG.; Fischer, MH.; Murray, WS.; Hill, RL., editors. Eye movements: A window on mind and brain. Elsevier; Amsterdam: 2007. p. 538-562.

Henderson, JM.; Hollingworth, A. Eye movements during scene viewing: An overview. In: Underwood, G., editor. Eye guidance in reading and scene perception. Elsevier Science Ltd; Oxford: 1998. p. 269-293.

Henderson JM, Weeks PA Jr. Hollingworth A. Effects of semantic consistency on eye movments during scene viewing. Journal of Experimental Psychology: Human Perception & Performance. 1999; 25:210–228.

Itti L, Koch C. A saliency-based search mechanism for overt and covert shifts of visual attention. Vision Research. 2000; 40:1489–1506. [PubMed: 10788654]

Johnston WA, Hawley KJ, Plewe SH, Elliott JMG, DeWitt MJ. Attention capture by novel stimuli. Journal of Experimental Psychology: General. 1990; 119:397–411. [PubMed: 2148574]

Jones BT, Bruce G, Livingstone S, Reed E. Alcohol-related attentional bias in problem drinkers with the flicker change blindness paradigm. Psychology of Addictive Behaviors. 2006; 20(2):171–177. [PubMed: 16784363]

Jones BT, Jones BC, Smith H, Copley N. A flicker paradigm for inducing change blindness reveals alcohol and cannabis information processing biases in social users. Addiction. 2003; 98(2):235–244. [PubMed: 12534429]

Koch C, Ullman S. Shifts in selective visual attention: Towards the underlying neural circuitry. Human Neurobiology. 1985; 4:219–227. [PubMed: 3836989]

LaBerge, D. Attentional processing. Harvard University Press; Cambridge: 1995.

Land M, Hayhoe M. In what ways do eye movements contribute to everyday activities. Vision Research. 2001; 41:3559–3566. [PubMed: 11718795]

Lee M, Shafran R. Information processing biases in eating disorders. Clinical Psychology Review. 2004; 24(2):215–238. [PubMed: 15081517]

Littel M, Franken IHA. The effects of prolonged abstinence on the processing of smoking cues: An ERP study among smokers, ex-smokers and never-smokers. Journal of Psychopharmacology. 2007; 21(8):873–882. [PubMed: 17984163]

Loftus GR, Mackworth NH. Cognitive determinants of fixation location during picture viewing. Journal of Experimental Psychology: Human Perception & Performance. 1978; 4:565–572. [PubMed: 722248]

Mackworth NH, Morandi AJ. The gaze selects informative details within pictures. Perception & Psychophysics. 1967; 2:547–552.

Mogg K, Bradley BP. Selective processing of smoking-related cues in smokers: manipulation of deprivation level and comparison of three measures of processing bias. Journal of Psychopharmacology. 2002; 16(4):385–392. [PubMed: 12503841]

Mogg K, Bradley BP. Attentional bias in Generalized Anxiety Disorder versus Depressive Disorder. Cognitive Therapy & Research. 2005; 29(1):29–45.

Mogg K, Bradley BP, Field M, De Houwer J. Eye movements to smoking-related pictures in smokers: relationship between attentional biases and implicit and explicit measures of stimulus valence. Addiction. 2003; 98:825–836. [PubMed: 12780371]

Mogg K, Field M, Bradley BP. Attentional and approach biases for smoking cues in smokers: An investigation of competing theoretical views of addiction. Psychopharmacology. 2005; 180(2): 333–341. [PubMed: 15696322]

Navalpakkam V, Itti L. Modeling the influence of task on attention. Vision Research. 2005; 45:205–231. [PubMed: 15581921]

NIH. Smoking measures - "mid-size" model baseline & follow-up measures. 2006. from http://www1.od.nih.gov/behaviorchange/measures/smoking.htm

Oliva A, Schyns P. Diagnostic colors mediate scene recognition. Cognitive Psychology. 2000; 41:176–210. [PubMed: 10968925]

Parkhurst D, Law K, Niebur E. Modeling the role of salience in the allocation of overt visual attention. Vision Research. 2002; 42:107–123. [PubMed: 11804636]

Shore TH, Tashchian A, Adams JS. Development and validation of a scale measuring attitudes toward smoking. The Journal of Social Psychology. 2000; 140(5):615–623. [PubMed: 11059207]

Torralba A, Oliva A, Castelhano MS, Henderson JM. Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. Psychological Review. 2006; 113(4):766–786. [PubMed: 17014302]

Treisman A. Features and objects: The Fourteenth Bartlett Memorial Lecture. Quarterly Journal of Experimental Psychology. 1988; 40A:201–237. [PubMed: 3406448]

Underwood G, Foulsham T. Visual saliency and semantic incongruency influence eye movements when inspecting pictures. Quarterly Journal of Experimental Psychology. 2006; 18:1931–1949.

Underwood G, Foulsham T, van Loon E, Humphreys L, Bloyce J. Eye movements during scene inspection: A test of the saliency map hypothesis. European Journal of Cognitive Psychology. 2006; 18(3):321–342.

Underwood, G.; Humphreys, L.; Cross, E. Congruency, saliency and gist in the inspection of objects in natural scenes. Eye movements: A window on mind and brain. van Gompel, RPG.; Fischer, MH.; Murray, WS.; Hill, RL., editors. Elsevier; Amsterdam: 2007. p. 564-579.

Walther D, Koch C. Modeling attention to salient proto-objects. Neural Networks. 2006; 19:1395–1407. [PubMed: 17098563]

Waters AJ, Feyerabend C. Determinants and effects of attentional bias in smokers. Psychology of Addictive Behaviors. 2000; 14(2):111–120. [PubMed: 10860110]

Waters AJ, Shiffman S, Sayette MA, Paty JA, Gwaltney CJ, Balabanis MH. Attentional bias predicts outcome in smoking cessation. Health Psychology. 2003; 22(4):378–387. [PubMed: 12940394]

Wolfe JM. Guided search 2.0: A revised model of visual search. Psychonomic Bulletin & Review. 1994; 1(2):202–238. [PubMed: 24203471]

Yarbus, AL. Eye movements and vision. Plenum; New York: 1967.

Yaxley RH, Zwaan RA. Attentional bias affects change detection. Psychonomic Bulletin and Review. 2005; 12(6):1106–1111. [PubMed: 16615336]

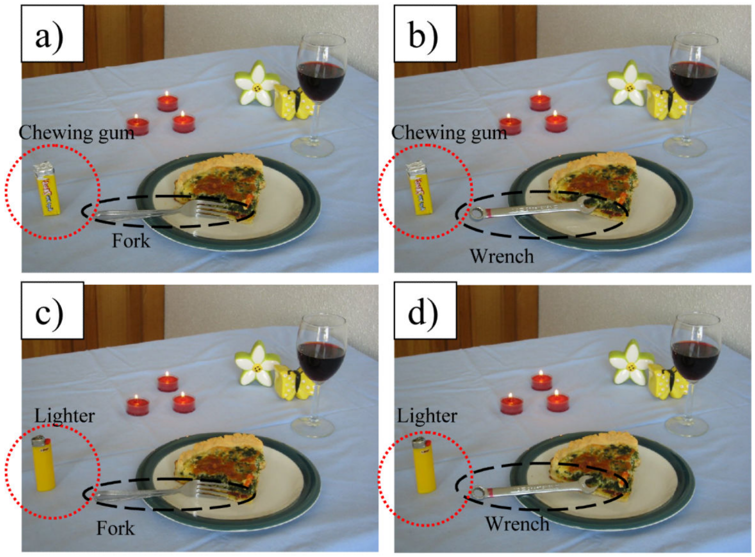

Figure 1.
Four Versions of a Dining Scene with Different Object Types: a) Consistent (Fork) / Neutral (Chewing Gum). b) Inconsistent (Wrench) / Neutral (Chewing Gum). c) Consistent (Fork) / Smoking-related (Lighter). d) Inconsistent (Wrench) / Smoking-related (Lighter).

Table 1

Mean (M) and Standard Deviation (SD) for all Target Objects

Target object	**Group**	***M (SD)***				
Context		**fixation count**	**dwell time (ms)**	**first fix. duration (ms)**	**number prior fix.**	**fixation probability**
Consistent						
Neutral	N	5.10 (0.98)	1566 (289)	230 (32.6)	7.66 (2.34)	0.92 (0.079)
	S	4.64 (1.13)	1631 (733)	241 (51.4)	7.70 (2.30)	0.92 (0.094)
Smoking	N	5.07 (0.98)	1507 (400)	233 (50.3)	7.18 (1.87)	0.91 (0.071)
	S	4.63 (1.13)	1485 (353)	238 (32.5)	7.63 (2.49)	0.92 (0.077)
Inconsistent						
Neutral	N	7.03 (1.60)	2096 (556)	249 (40.9)	6.30 (2.07)	0.99 (0.032)
	S	6.99 (2.50)	2267 (628)	264 (49.1)	5.89 (1.21)	0.96 (0.050)
Smoking	N	7.20 (1.44)	2236 (462)	236 (34.7)	6.07 (1.60)	0.97 (0.058)
	S	7.01 (1.97)	2301 (638)	253 (39.5)	6.31 (1.23)	0.98 (0.055)
Neutral						
Consistent	N	4.83 (1.67)	1467 (502)	260 (49.1)	7.31 (1.59)	0.95 (0.061)
	S	4.16 (1.00)	1303 (369)	261 (44.2)	6.77 (1.34)	0.95 (0.017)
Inconsistent	N	4.47 (1.41)	1366 (355)	252 (48.5)	7.91 (2.21)	0.91 (0.129)
	S	3.98 (1.07)	1212 (286)	253 (33.5)	6.70 (1.59)	0.93 (0.024)
Smoking						
Consistent	N	4.78 (1.38)	1481 (381)	279 (50.4)	7.38 (1.87)	0.94 (0.084)
	S	5.48 (1.93)	1769 (551)	280 (69.2)	5.69 (1.05)	0.98 (0.016)
Inconsistent	N	4.74 (1.27)	1421 (309)	262 (60.8)	7.92 (2.19)	0.93 (0.023)
	S	4.66 (1.12)	1519 (449)	291 (70.3)	6.84 (1.39)	0.97 (0.024)

Note. Group N = non-smokers (N = 19); group S = smokers (N = 18)