Generations of scientists and philosophers have struggled with the mystery of the possible functions of conscious awareness. Although some scholars deny that consciousness has any functional role (e.g., Velmans, 2009), others see it as a necessary condition for decision making, action selection, and self-deliberation (Mandler, 2002); appropriate performance in unusual tasks and situations (Dehaene & Naccache, 2001); error identification and correction (Posner, 1998); and planning (Crick & Koch, 2003). Common to all the theories that assign functional significance to awareness is the notion that information integration is one of its fundamental features: Whereas they hold that conscious awareness is not required for low-level perceptual binding (e.g., of object parts), they postulate that consciousness is necessary for rapidly joining together perceptual and conceptual data from diverse sources to create a unified and coherent scene or idea (e.g., Tononi & Edelman, 1998).

Terms used to describe the neural correlate of consciousness, such as “global workspace” (Baars, 2005) and “coalitions of neurons” (Crick & Koch, 2003), clearly convey the same notion concerning the role of consciousness in integration. According to the awareness-as-integration views mentioned in the previous paragraph, awareness can be described as the ability to establish specific relationships between representational items, thereby enabling the formation of a structured mental representation (Engel, Fries, Konig, Brecht, & Singer, 1999; Goodale, 2004). These views assign to awareness a constructive function that is indispensable for perceiving and comprehending the meaning of scenes (Mandler, 2002; Marcel, 1983). This idea seems to be consistent with the fact that, despite the distributed nature of perceptual coding, the contents of awareness are phenomenologically perceived as unified, coherent wholes rather than as a mosaic of constituent parts. Thus, most theories predict that a conceptual relationship between objects is not established, and may be impossible, under unconscious perception.

In this study, we examined this prediction, which has not been directly tested before. Indeed, previous attempts at...
demonstrating semantic unconscious processing have been limited to stimuli consisting of a single word or object, and there is currently no evidence for integration among multiple components of a visual scene that is not consciously perceived (for a review, see Kouider & Dehaene, 2007). Therefore, we looked for integration of a meaningful object into a scene when neither the object nor the scene was consciously perceived. Unawareness was achieved by using the continuous flash suppression (CFS) method (Tsuchiya & Koch, 2005). In CFS, distinct color images (Mondrians) presented successively at approximately 10 Hz to one eye can reliably suppress the conscious awareness of an image presented to the other eye for a relatively long duration. Previous CFS studies have demonstrated high-level unconscious processing of single faces, words, and facial expressions; they have shown that upright faces and recognizable words can break suppression faster than inverted faces and unrecognizable words, respectively (Jiang, Costello, & He, 2007), and that fearful expressions emerge into awareness faster than neutral ones (Yang, Zald, & Blake, 2007). The logic underlying the interpretation of these studies is that if the effectiveness of suppression differs between images that differ on some dimension, this dimension must be processed while the images are suppressed; otherwise, one image could not break suppression more quickly than the other.

Within this framework, we used CFS to suppress awareness of scenes in which a critical object was either congruent or incongruent with the overall context (e.g., a man drinking from a glass vs. a man “drinking” from a hairbrush). Low-level features, including contrast, brightness, chromaticity, and spatial frequency, as well as the perceptual saliency of the objects, were matched across congruent and incongruent scenes (see Low-Level Features Evaluation in the Supplemental Material available online), so that the only difference between these scenes was the contextual relation between the critical object and its background scene. The scenes appeared on the left or right of fixation, and subjects were required to press one of two buttons as soon as they detected the hemifield in which the image appeared—a task that does not require any semantic processing (Jiang et al., 2007). We compared the time it took subjects to respond to congruent scenes and to incongruent scenes. We reasoned that if the integration between an object and its background scene can be achieved even when they are not consciously perceived, the duration of suppression should differ between the two kinds of stimuli (i.e., a congruency effect).

Method

Subjects

Subjects were 18 healthy volunteers (12 females and 6 males), students of the Hebrew University of Jerusalem (age 19–30 years, $M = 23$ years). All subjects had reportedly normal or corrected-to-normal sight and no psychiatric or neurological history. They participated in the study for payment (about $5 per hour). The experiment was approved by the ethics committee of the Department of Psychology at The Hebrew University of Jerusalem, and informed consent was obtained after the experimental procedures were explained to the subjects. Three additional subjects were excluded from analysis because they had difficulty reaching fusion of the images between the two eyes during the calibration phase at the beginning of the trials.

Stimuli and procedure

Using a mirror stereoscope, subjects viewed test scenes ($2.86\times 2.03\) with one eye while simultaneously viewing suppressors ($5.26\times 5.26\), replaced every 100 ms, with the other eye. MATLAB Psychophysics Toolbox (Brainard, 1997) was used to present stimuli on a 17-in. CRT monitor with a 100-Hz refresh rate. The stereoscope was connected to a chin rest and was positioned in front of the monitor.

The test scenes were color pictures taken from Internet sources. Each showed a human taking action involving an object (e.g., a man playing a violin, a woman using a microscope). An incongruent version of each scene was created by replacing the original object of the action with another, unrelated object (e.g., basketball players holding a watermelon, instead of a basketball; see Fig. 1c). To equate the amount of digital manipulation in the congruent and incongruent images, we replaced the object in congruent scenes by another exemplar of the same object category (e.g., the basketball in the original scene was replaced by a basketball copied from another image). Brightness and contrast were equated using Photoshop. To rule out other systematic differences between the scene types, we compared the chromaticity and saliency maps of the congruent and incongruent images using dedicated algorithms (see Low-Level Features Evaluation in the Supplemental Material). All the processed pictures were pretested in a separate experiment, in which subjects ($N = 24$) viewed each scene for 1 s and then answered the question, “How unusual was the picture, in your opinion?” Responses were made on a scale ranging from 0 (not at all) to 4 (very abnormal). Twenty-eight pairs in which either the congruent image was rated higher than 1 by more than 30% of the subjects or the incongruent image was rated lower than 3 by more than 30% of the subjects were excluded. The final set included 42 pairs.

The suppressors were Mondrians, random amalgams of partly overlapping rectangles of varying sizes and colors. All stimuli (Mondrians, congruent scenes, and incongruent scenes) were surrounded by rectangular borders ($5.86\times 5.86\) that served to promote stable binocular eye alignment.

The experimental session included 80 CFS trials, divided into two blocks. In each block, congruent scenes were presented in half of the trials, and incongruent scenes were presented in the other half. Congruent and incongruent versions of the same scene were always presented in different blocks. The two scene types (congruent vs. incongruent) were
intermixed, with the constraint that the same type was never presented in 4 or more consecutive trials.

Each trial started with a stereoscope calibration phase (see Calibration Phase—Procedure in the Supplemental Material). Next, a Mondrian was presented to the subject’s dominant eye at full contrast, and the test scene was presented to the other eye, to either the right or the left of fixation. The contrast of the test scene was ramped up gradually from 0% to 100%, with a 10% increment every 100 ms. After 1 s, when the test scene had reached full contrast, the contrast of the Mondrian decreased linearly at a rate of 2% every 100 ms for the next 5,100 ms. In each trial of the control condition (b), the test scene was blended into the dynamic noise pattern of the Mondrian and presented binocularly; contrast of the scene was ramped up at a rate of 2.5% every 100 ms. In both conditions, each scene was shown in both an incongruent version and a congruent version in separate trials. The example scenes shown here (c) depict a woman putting either food or a chessboard in the oven, a boy holding a bow and either an arrow or a tennis racket, and two athletes playing basketball with either a ball or a watermelon.

The control condition included 80 trials, divided into two blocks. In each block, half of the trials were congruent, and the other half were incongruent. Scene types were intermixed. The congruent and incongruent images from each pair were presented in different blocks. In the control condition, the test scenes used in the CFS trials were blended into the Mondrians, and the blended images were presented to the two eyes (i.e., no rivalry; see Fig. 1b). The contrast of the test scenes in the stimuli started at 0% and was ramped up gradually at a rate of 2.5% every 100 ms, so that the scenes reached full contrast after 4 s. Subjects were asked to perform the same task as in the CFS condition (i.e., to press a key to indicate the location of the test stimulus as soon as it was detected). The cardinal difference between the CFS and control conditions was that in the CFS condition, the test scenes were fully exposed, albeit without...
awareness, whereas in the control condition, the scenes were only gradually exposed as the signal-to-noise ratio increased. If partial awareness during rivalry caused congruency effects to occur in the CFS condition, then such effects would also emerge in the control condition, because partial awareness was equally likely to occur in the two conditions. By contrast, finding congruency effects in the CFS condition only would allow us to reject the partial-awareness account and to conclude that the congruency effects resulted from the availability of the suppressed information during rivalry.

Finally, an evaluation task was included after the CFS and control conditions were completed. In this task, all the stimuli were presented binocularly without any noise (dynamic or static). Subjects were asked to report whether they thought that each presented scene was unusual or not (i.e., incongruent or congruent) by pressing either the right or the left key of a response box. This phase was included to validate that subjects could detect the context violations when perceiving the scenes through the stereoscope and to allow us to omit from analysis scenes that an individual subject did not classify correctly (see Results).

At the beginning of each stage of the experiment (CFS block, control block, and evaluation task), subjects performed two training trials to ensure that they were able to perform the task for that block and understood the classification requirements. The training trials included one congruent trial and one incongruent trial (the order of the trials was counterbalanced across subjects), and scenes that were presented during training were not presented during the experimental session. The training trials were identical to the corresponding experimental trials except for the fact that feedback on the response (i.e., the location of the test scene or congruency) was given only during training.

Results

Congruency ratings

Subjects evaluated most congruent and incongruent scenes correctly in the evaluation task (congruent scenes: $M = .87, SD = .08$; incongruent scenes: $M = .87, SD = .07$). Scenes that a given subject did not classify correctly were excluded from all further analyses for that subject. Thus, for each subject, only scenes that were correctly rated as congruent or incongruent were included in the analyses. Note that the evaluation stage always followed the CFS and control blocks.

Accuracy

In the CFS condition, subjects were very accurate in reporting the location of the scene (left vs. right), and accuracy did not differ between congruent scenes ($M = .97, SD = .03$) and incongruent scenes ($M = .96, SD = .04$), $t(17) = 1.94, p > .05$. In the control condition, subjects were even more accurate, and again, accuracy did not differ between congruent scenes ($M = .99, SD = .02$) and incongruent scenes ($M = 1.00, SD = .01$), $t(17) = 1.41, p > .05$.

Time to emerge into awareness

In the CFS condition, incongruent scenes emerged into awareness earlier ($M = 2.50 s, SD = 0.60 s$) than congruent scenes did ($M = 2.64 s, SD = 0.62$), $t(17) = 2.82, p < .02$ (two-tailed, paired $t$ test; Fig. 2). Because the congruency between an object and its background affected suppression time, we may conclude that the objects and backgrounds must have been integrated while the scenes that they composed were blocked from conscious awareness.

In contrast with the results for the CFS condition, results for the control condition showed no difference in suppression time between congruent ($M = 1.99 s, SD = 0.41$) and incongruent ($M = 1.97 s, SD = 0.38$) scenes, $t(17) = 0.93, p > .05$ (Fig. 2). A joint analysis of the two conditions in a two-way repeated measures ANOVA showed an interaction between condition (CFS, control) and scene type (congruent, incongruent), $F(1, 17) = 6.56, p < .05$, in addition to significant main effects. Because effects of partial awareness were as likely to have been present in the CFS condition as in the control condition, it is improbable that the congruency effect found in the CFS condition resulted from partial-awareness effects.

Next, we addressed the possibility that a congruency effect was found in the CFS condition but not in the control condition because detection times were longer overall in the CFS condition. We removed from analysis the 5 subjects for whom mean detection times in the CFS condition exceeded the highest mean detection time observed in the control condition and found that the congruency effect remained significant, $t(12) = 2.32, p < .05$ (two-tailed). Thus, the incongruency advantage found in the CFS condition likely reflects perceptual availability of the suppressed information during the CFS (but not control) trials.

Discussion

Our results provide clear evidence of conceptual integration between an object and its background without conscious awareness of either. Although subjects did not consciously perceive the congruent and incongruent scenes they were viewing, they not only processed the displayed object and its background, but also attempted to integrate the two into a meaningful and coherent scene. By contrasting the results from the CFS condition with those from the control condition, we invalidated an alternative interpretation of the results (i.e., that subjects were partially aware of the stimuli, and that recognition speeds or detection criteria differed between the two scene types).

An important question for future research is whether integration without awareness is unique to manipulable objects and their contexts, or whether our findings can be generalized to other classes of objects. Initial evidence suggests that tools might be more intensively processed than other objects under
putatively because processing tools involves the dorsal visual system, which mediates functional ("how") knowledge (Almeida, Mahon, Nakayama, & Caramazza, 2008; Fang & He, 2005; Goodale & Milner, 1992).

A second question involves the processes conferring context congruency. Processing the gist of a scene may activate abstract, semantic knowledge that could facilitate recognition of (or instigate predictions about) scene-relevant objects in a top-down manner (e.g., Bar, 2004; Mudrik, Lamy, & Deouell, 2010). Alternatively, representations of frequently co-occurring objects may more readily interact with one another than representations of unrelated objects do because they draw on learned probabilities, and do not need to involve top-down predictions or abstract knowledge. This distinction between abstract knowledge and learned probability of occurrence is sometimes impossible to make, especially in the case of natural scenes. However, many of the scenes we used included just the user and the object (e.g., players holding a basketball or a watermelon, a man drinking from a glass or a hairbrush), without any other congruent or incongruent objects that could interact with each other. In such cases, only semantic, or functional, knowledge was available. Regardless of the specific mechanism underlying the context-congruity effect, our results suggest that the involved integrative processes do not require conscious awareness.

The findings reported here thus expand the known limits of unconscious processing. Previous studies demonstrated processing without awareness, for example, in lexical and semantic judgments (e.g., Lamy, Mudrik, & Deouell, 2008), odd/even judgments (Bodner & Dypvik, 2005), emotional-valence identification (e.g., de Gelder, Pourtois, & Weiskrantz, 2002), meaning assessment (Jiang et al., 2007), and motor preparation for action following unconscious processing (Dehaene et al., 2001; Jiang, Costello, Fang, Huang, & He, 2006), yet all of the tasks employed in these studies pertained to single items: a word, a digit, or an object. Thus, previous findings do not speak to the question of high-level integration between objects making up a meaningful scene. By showing that awareness is not necessary for grasping the semantic relations between the constituents of a visual scene and for establishing structured mental representations, our results widen the known boundaries of unconscious processing to territories previously assigned exclusively to the domain of conscious awareness and challenge prominent theories of awareness that put major emphasis on integration as a defining feature of consciousness (e.g., Goodale, 2004; Mandler, 1984; Marcel, 1983).

According to these theories, consciousness serves as the primary agent that facilitates widespread access between massively distributed sets of specialized neural networks (global workspace theory; see Baars, 1988, 2005). In other words, without awareness, processing in discrete nodes or small networks proceeds in parallel, without integration and thus without conflict. By showing that integration of objects and their background scenes can be achieved without awareness of either, our findings undermine the traditional view that limits unconscious processes to simple operations not involving

![Fig. 2. Suppression durations (detection times) for congruent and incongruent scenes in the continuous flash suppression (CFS) condition (left panel) and control condition (right panel). Each triangle represents the average suppression duration for the two scene types for an individual subject (congruent scenes: x-coordinate; incongruent scenes: y-coordinate). The diagonal line represents equal suppression durations for congruent and incongruent scenes; data points below the diagonal indicate shorter suppression times for incongruent scenes.](image_url)
complex meaning, high-level inferences, executive control, or intention (e.g., Baars, 1988; Cheesman & Merikle, 1984). Thus, our findings are in line with the results of recent studies demonstrating unconscious activation of a cognitive-control system (Lau & Passingham, 2007), implicit goal pursuit (Hassin, Bargh, & Zimerman, 2009), and unconscious flexibility (Stapel & Koomen, 2006).

If integration can occur during unconscious processing, what role is left for conscious awareness? At first glance, our findings might be considered to support epiphenomenal accounts (e.g., Velmans, 2009), which leave no functional role to awareness. Such accounts hold that all cognitive and perceptual processes are performed unconsciously, but that some are accompanied by sensations (i.e., qualia, the qualitative character of experience) that have no influence on the chain of events taking place in the brain. These accounts further posit that consciousness is nothing but a by-product of the brain’s complexity, much as the whistle of a steam locomotive is the result of the physical processes taking place in the engine, but has no effect on the train (Huxley, 1874/1898). Thus, according to these accounts, there cannot be any functional difference between conscious and unconscious processes. They differ only in subjects’ ability to report the accompanying sensations (to other people or to themselves).

However, our findings may actually highlight a crucial role of awareness in perception and scene interpretation. In our study, incongruent scenes were faster than congruent scenes to break suppression1 and emerge into awareness, presumably because attempts at integrating the objects and backgrounds in these scenes yielded a conceptual difficulty. The process of recognition may involve rapid formation of initial guesses or predictions, which interact with gradual accumulation of data (Bar, 2004, 2007). Possibly, then, although awareness is not needed for object-background integration per se, it may become necessary when this integration yields a conceptual conflict, which might signal a potentially hazardous or fortuitous situation. In other words, perceptual flexibility in the face of novelty may call for awareness (Dehaene & Naccache, 2001; Searle, 1992; van Gulick, 1989). Taken together with previous research, our findings suggest that the “zombie within” (Koch & Crick, 2001), that is, the unconscious processes underlying perception, behavior, and cognition, may be much more sophisticated than was previously thought. Nevertheless, the sophistication of the unconscious may not be all encompassing; every once in a while, in novel situations or when things go wrong, conscious awareness may be needed to set things straight.

Acknowledgments
We thank Liron Levin for help in programming, Tal Golan for important help in image processing and analysis, and Ran Hassin and Noam Sobel for helpful discussion and comments.

Declaration of Conflicting Interests
The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Supplemental Material
Additional supporting information may be found at http://pss.sagepub.com/content/by/supplemental-data

Note
1. Note that we are not saying that incongruent scenes are processed more quickly than congruent ones; rather, we are arguing only that incongruent scenes breach suppression and emerge into awareness more quickly. Previous studies (e.g., Chun & Jiang, 1998; Davenport & Potter, 2004) suggest that congruent objects are recognized more quickly than incongruent ones, a finding that is probably attributable to utilization of top-down predictions during recognition (Biederman, Mezzanotte, & Rabinowitz, 1982).

References


