Article

Semantic priming modulates the strength and direction of the Kanizsa illusion

Check for updates

Nataly Davidson Litvak[1,4], Amir Tal[1,2,4] & Liad Mudrik [1,2,3] ✉

Visual illusions are considered key examples for cognitive impenetrability, as they are held not to be affected by non-perceptual processes. We revisit this claim in five experiments ($N = 1148$; four preregistered) focused on the Kanizsa illusion, where a nonexistent shape is experienced within illusory contours. Pac-Man-like shapes inducing the illusion were presented after primes that were either semantically related to the Pac-Man game or not. We hypothesized that semantic primes would promote interpreting the shapes as individual Pac-Man characters, thus biasing participants away from the holistic Kanizsa illusion. Indeed, we found that the Kanizsa shape was detected less when participants were primed with Pac-Man-related stimuli. We then also demonstrated the opposite effect: a prime indexing the illusory shape ("Triangle") enhanced the probability of seeing the illusion. Together, our results suggest that semantic priming can both reduce and increase the probability of experiencing the Kanizsa illusion, thus supporting claims of cognitive penetrability.

"Cognitive penetrability" is the notion that higher mental states can affect perceptual processing[1–3]. Under this view, perception integrates both bottom-up (i.e., incoming input from the sensory stimuli) and top-down, non-perceptual factors (e.g., knowledge, motivations, beliefs, emotions, linguistic representations). Arguably, each percept thus combines the actual sensory properties of the stimulus with a variety of cognitive factors that affect the interpretation of these properties[4,5]. This view seems to align well with more recent 'predictive coding' accounts[6,7], depicting the brain as a non-passive updating system that generates predictions about the world, based on inner schemes[8,9]. These predictions are continuously updated through feedforward and feedback connections[7,10] and are reflected in processes of neural re-organization and synapse strength modulation[11,12]. And so, some claim that we continuously "hallucinate reality" in a controlled manner, rather than accurately reflect the external world as it is[13]. Indeed, many studies demonstrated different types of top-down effects on perception (for reviews and discussion see refs. 1,14), held to be more pronounced when the physical stimulus is degraded or ambiguous (e.g., refs. 15,16). Proponents of cognitive penetrability further claim that such a penetrable system is more efficient and evolutionarily desirable, as it considers all available information and aims at the most relevant interpretation of the world, even at the risk of not always being a veridical one[1].

Yet, some still strongly deny that perception can be cognitively penetrated[17], supporting a modular view according to which perception is encapsulated from cognition[18]. Under this view, there is a foundational component of visual processing—"early vision"[19]—that is determined solely by physical properties of the stimulus. Accordingly, any effects of expectations, desires, beliefs, etc. occur either before that component, in the form of pre-perceptual attention allocation, or after it, via post-perceptual processes like inference, evaluation and selection[19]. Proponents of cognitive impenetrability accordingly reject demonstrations of cognitive penetrability, claiming that they either misattribute pre-perceptual or post-perceptual effects to perception[19], or that they suffer from methodological limitations[17,20]. Firestone and Scholl[17] criticize current literature for taking an overly confirmatory (as opposed to disconfirmatory) approach, and list five potential confounds that, arguably, can explain the results without assuming cognitive penetrability: judgment effects, demand and response biases, low-level differences between the stimuli in the different conditions, peripheral attentional effects and memory effects. To give an example for this line of criticism, one study reported that wearing a heavy backpack biases participants towards judging hills as steeper[21]. However, this effect went away when the experimenters offered a plausible cover story for why the backpack was required in the experiment[22]. Apparently, when no such story was given, participants were cognizant of the study manipulation to some degree and adjusted their responses accordingly to please the researchers. In another example, faces of black people, matched in mean luminance with those of white people, were nevertheless judged as seeming darker[23]. Yet when taking a disconfirmatory approach, blurring these faces such that their race was no longer detectable, the black faces were still rated as darker, suggesting that other low-level features than luminance still evoked lighting differences[20]. Based on such examples, and many others, Firestone and Scholl argued that

[1]School of Psychological Sciences, Tel Aviv University, Tel Aviv, Israel. [2]Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel. [3]Brain, Mind, and Consciousness Program, Canadian Institute for Advanced Research (CIFAR), Toronto, Canada. [4]These authors contributed equally: Nataly Davidson Litvak, Amir Tal. ✉e-mail: mudrikli@tauex.tau.ac.il

current findings cannot support cognitive penetrability[17] (though see refs. 24–26).

The debate on cognitive penetrability has far reaching implications, and it is hence the locus of an ongoing philosophical and psychological controversy[3,27]. Within this debate, visual illusions hold a special role. In visual illusions, perception is inaccurate, typically due to a misleading interpretation of the physical properties of the stimulus. So, for instance, the same exact color may appear distinctly different when it is located within a shaded or an unshaded area of a stimulus[28]. Despite this seemingly higher-level effect on lower-level visual processing (argued by some to support penetrability;[3,10]), visual illusions are notoriously resilient and impervious to cognition. So, for instance, the aforementioned illusion is not abolished by knowing that the colors are the same, or even after experiencing that this is the case (e.g., when the two appear in isolation without the background and cues of shading). Visual illusions therefore seem to demonstrate that perception is encapsulated from cognition, as "knowing about them does not make them disappear"[19], and so are typically presented as a definitive case in point in favor of the impenetrability approach[29,30], or as "the most striking evidence in favor of cognitive impenetrability"[31]. Thus, illusions pose an especially interesting test case to the claim that cognition can affect perception. Can they be altered by a cognitive state? If so, this would lend strong support to the cognitive penetrability hypothesis.

Thus far, such alterations have scarcely been demonstrated. Indirect evidence can be drawn from works showing that the perception of some visual illusions is culturally modulated. For instance, Toda individuals from the Indian subcontinent were less susceptible to the Müller-Lyer illusion[32] than English individuals[33], and Himba individuals from Namibia were less susceptible to the Ebbinghaus illusion[34] than English participants[35]. These differences seem to indicate culturally-dependent development of perceptual biases[36], which have been tied to urbanicity[37]. As such, however, they can nevertheless align with a cognitive impenetrability account, as perceptual biases can be considered within the compartment of early vision and hence, albeit sensitive to external stimuli over the course of development, still operate in isolation from cognition[19]. Another study reported that the neural response to the Müller-Lyer illusion in the intraparietal sulcus was modulated by the task, such that it was stronger when participants were asked to judge the length of the line[38]. Notably though, the results were limited to the neural response: since the other task did not involve length estimation, there is no way to know if the illusion was indeed experienced as stronger. In addition, proponents of impenetrability could still ascribe the effect to attention rather than perception per se (for earlier work describing the effect of attention on the perceived illusion, see ref. 39).

Stronger evidence for the potential effect of cognition on visual illusions comes from three studies that directly investigated top-down semantic influences on visual illusions. Two of them focused on the Ternus illusion of apparent motion [40]. In this illusion, two displays containing a similar array of objects (e.g., three equally spaced identical circles) alternate. The displays differ in the location of the circles: in the second array, all stimuli are shifted to the right, such that the location of the most extreme stimulus to the left is empty, and instead the most extreme location on the right is occupied. This typically elicits a visual illusion of movement, with its nature depending on the speed of the alternation: Quick shifts between the two displays cause an illusion of element motion, as the extreme element seems to jump from side to side while the others remain in place. Slower shifts elicit an illusion of group motion, as the entire array seems to move in unison, maintaining the relative position of the elements.

In both the aforementioned studies, the Ternus illusion was affected by semantic information. In one study, the type of motion that was perceived depended on whether the "moving" elements (two ellipses) were interpreted as a person's feet or as wheels of a car. Seeing them as feet enhanced element motion while seeing them as wheels enhanced group motion, in accordance with real-life knowledge that feet motion entails one foot moving across the other, while wheel motion entails joint movement[41]. In a second study, realistic pictures of frogs were used as the elements of the illusion, and these were either facing the direction of the illusory movement or with their backs to it. The latter case elicited greater illusion of element over group motion, arguably due to real-life knowledge that frogs leap forward, and so a leap over a frog was more plausible when the leaper was facing the direction of the leap[42].

Finally, the third paper focused on the uniformity illusion, where fixating at the center of a stimulus in which the central area differs from its peripheral areas in some dimension (e.g., blurriness, color, contrast) leads to the disappearance of this difference. The entire stimulus then seems uniform, with the peripheral area appearing like the central one[43]. The researchers assumed that illusions such as this one, that rely on peripheral vision, are especially susceptible to demonstrations of cognitive penetrability, because bottom-up peripheral information is noisy and thus more prone to be overridden by top-down perceptual priors [29]. They therefore tested whether they could influence this illusion by manipulating the weighing of bottom-up sensory evidence vs. top-down perceptual priors, using affective state. Indeed, they found that inducing negative mood using music reduced the chance of seeing the illusion and prolonged its onset when it was perceived, in accordance with the notion that negative mood biases perception towards minimal discrepancies between bottom-up and top-down information (i.e., minimizing prediction errors)[44].
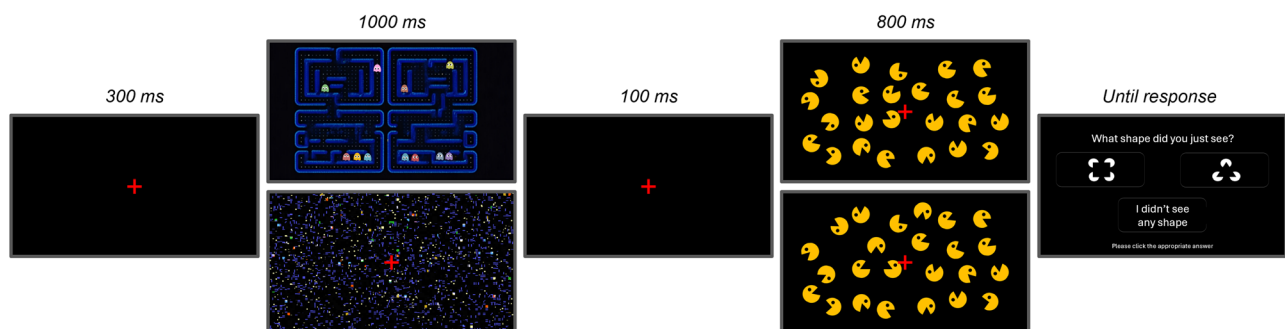

Fig. 1 | Trial sequence. Each trial began with a fixation cross on top of a black screen (300 ms). It was then followed by either a semantic or a non-semantic prime display (1000 ms). Here, prime displays from Experiment 1 are shown: A Pac-Man gameboard (top) was the semantic prime, and a scrambled version of that gameboard (bottom) was the non-semantic prime. Next, an additional black fixation screen was shown (100 ms), followed by either a no-Kanizsa display (containing no Kanizsa illusion; top) or a Kanizsa display (containing one Kanizsa triangle to the left of fixation; bottom; 800 ms). Finally, an answer screen appeared with the question "What shape did you see?". This screen was shown until participants selected one of the following options: a Kanizsa square, a Kanizsa triangle (locations counterbalanced across participants), or "I didn't see any shape". The answer screen was in Hebrew in the experiment; it is shown here in English for clarity. Due to copyright issues, the image of the Pac-Man gameboard shown in this figure is for illustration and not the one we have used in the experiment. This image was created by drawing over an image obtained from Shutterstock.

Notably though, the first two studies only demonstrated a modulation of the illusion based on semantic information, biasing perception towards group or element motion. The third went one step further, showing a decrease in the probability to see the illusion following negative mood induction, but the effect was relatively small (3.6%). And, to date, no study has shown that participants might even not see an illusion altogether due to semantic influences.

In the current study, we accordingly set to examine if a well-known and highly robust visual illusion—the Kanizsa illusion[45]—could be abolished in some participants following semantic priming. In the Kanizsa illusion, contextual inducers are organized in a way that evokes the perception of illusory shape contours, though no explicit lines or enclosed spaces appear. We used a specific variant of the illusion, where the contextual inducers resemble the cartoon character from the video game "Pac-Man" (see Fig. 1). We presented either a Pac-Man-related prime or a neutral prime prior to the Kanizsa display, under the assumption that the former would bias participants towards perceiving the sliced circles as individual Pac-Man figures, thereby reducing the chances of them being grouped together to form the contour of a Kanizsa triangle. From a figure-ground perspective[46], for the Kanizsa illusion to occur, the observer should perceive the inducers as the "ground" components, allowing them to form an illusory shape (which becomes the figure)[47]. We hypothesized that priming the Pac-Man context would activate related representations (e.g., Pac-Man figures), prompting the identification of the inducers as these figures, as opposed to the ground components of the display. This accords with previous studies showing that perceiving the gist of a scene (here, the Pac-Man board, or a verbal label serving as prime) automatically activates representations of objects that are likely to appear in it[48]. This, in turn, increases the chances of upcoming information to be interpreted as matching one of the activated templates[49]. This hypothesis was tested in a series of five preregistered experiments, examining the potency of semantically related images and words to reduce – and also facilitate – perception of the Kanizsa illusion.

Methods
Open science practices
Four out of the five experiments were preregistered (Exp. 1 on Aug 2019, Exp.2 on Aug 2019, Exp. 3 on Jan 2020, and Exp. 5 on Nov 2020; Exp. 4 was also meant to be preregistered, and a preregistration file had been prepared prior to running the experiment, but the file was not uploaded to OSF due to a technical error. It has now been added to the OSF page). The preregistration files, materials, data and analysis codes are all available on the Open Science Framework: https://osf.io/wmak2/.

Participants
Overall, 1148 participants were included in this study over the five experiments. For experiments 1–4, 200 participants were required to provide power higher than 0.99, considering a medium effect size of 0.25 and a standard 0.05 alpha (calculated using the G*Power software). This was therefore the sample size used in the four experiments (Experiment 1: 88 female participants, 112 male participants; age 18–35, $M = 25.9$, SD = 4.4. Experiment 2: 106 female participants, 94 male participants, age 18-35, $M = 25.2$, SD = 4.2. Experiment 3: 97 female participants, 103 male participants; age 18–35, $M = 24.6$, SD = 3.6. Experiment 4: 78 female participants, 122 male participants; age 18–35, $M = 25.4$, SD = 3.6). Twenty-one additional participants were excluded from analysis due to extreme RT (seven from Experiment 1, two from Experiment 2, four from Experiment 3 and eight from Experiment 4; see *Exclusion Criteria* below), and one (from Experiment 4) due to technical issues during the experiment.

For Experiment 5, 348 participants were required, as it included three experimental groups. Here, power analysis was conducted using the R "powerAnalysis" package, with the "power.chisq" function and an effect size of 0.25, power of 0.99, 2 degrees of freedom and an alpha of 0.05. The sample size accordingly included 348 participants (213 female participants, 133 male participants, 2 participants selected "other" for sex; age 18–35,

$M = 28.9$, SD = 4.6). Six additional participants were excluded due to extreme RT.

Participants in experiments 1–4 were all Hebrew speakers, with normal or corrected-to-normal vision, approached at public places. Participation was on a voluntary basis (no payment was offered). Due to COVID-19, participants in experiment 5 were recruited remotely, using Amazon Mechanical Turk (Mturk), for a payment of 0.4 USD. These participants were all English speakers, who met the following criteria: HIT approval rate greater than 98, number of HITs approved greater than 500, and age between 18 and 35. They were required to perform the experiment on the Qualtrics platform on a PC. For all experiments, participants signed an informed consent, as approved by the Tel Aviv University ethics-committee.

Condition X group allocation was pseudo-random: it was pre-assigned internally to each participant number (unbeknownst to the experimenter), making sure that there would be a roughly equal number of participants in each group.

Exclusion criteria
Mean RT was calculated for each group (irrespective of answer). Participants with RTs that diverged from the average RT of their group by three standard deviations (+3/−3) or more were excluded from further analysis, as reported above.

Stimuli
The test displays contained 24 sliced yellow circles ("Pac-Man" characters; Fig. 1). In the triangle Kanizsa display, three of them were arranged to create an illusionary triangle[45]. The decision to embed the Kanizsa display within an array of shapes, instead of using the typical, isolated and foveally presented display, was motivated by two reasons. First, we assumed that such a display would facilitate potential top-down effects, given that the bottom-up stimulation is noisier[29]. Second, we wanted to avoid a ceiling effect, where all participants would easily see the shape. These considerations further motivated the decision to present the stimulus for 800 ms only (see further detailed below). In the No-Kanizsa display, the same sliced circles were rotated in a way that creates no illusory Kanizsa shape. To avoid low-level differences between the displays, we equated the number of pixels of each color using a MATLAB code. Several pilot experiments were conducted to ensure that the Kanizsa display is indeed effective (i.e., the shape is detected by at least 75% of the viewers). Because the study was conducted outside the lab while participants were standing, their head was not fixed and their viewing distance somewhat varied around ~50 cm.

Based on the pilot experiments, the illusory triangle's sides were set to be 3.9 × 3.4 × 3.4 cm (4.5° × 3.9° × 3.9° visual angles). It was positioned a bit over the vertical center of the screen (8.6° from the bottom of the screen, 7.2° from the top) and in the left of the horizontal center (11.1° from the left screen border, 17.2° from the right). Other sliced circles were at least 1 cm (1.1° visual angles) distant from the three Kanizsa inducing circles. The background of the pictures and instructions was black. The instructions were written in white Hebrew letters, in David font.

The Pac-Man prime in Experiment 1 was a screenshot taken from the "Pac-Man" video game, presenting a Pac-Man gameboard maze (see again Fig. 1 for illustration). The Pac-Man character itself was removed from the picture to prime only the general context of the game and not the character itself. The non-semantic "scrambled" prime was a scrambled version of the Pac-Man display picture, created using Adobe Photoshop CC2018's "Scramble" filter. Tile parameters were 0.52 cm height and 0.52 cm width (20 px × 20 px or 0.6° × 0.6° visual angles).

The gift box prime in Experiment 2 was created using elements from the Pac-Man prime, aiming to obtain similar colors and style (cartoon; see Supplementary Fig. 1). The number of color pixels in both primes was equated using a MATLAB code, to reduce low-level differences between the two primes. The display's background was black and the gift box (which was 3D-shaped, with 10.9° x 14.3° x 10.3° dimensions) appeared in the middle of the screen. A pre-test was conducted to ensure that the new display did not remind participants of the Pac-Man videogame. The non-semantic

scrambled display was a scrambled version of the gift box prime, created in the same manner as in Experiment 1.

The prime words in Experiments 3 and 4 were the Hebrew word for Pac-Man (פקמן) and a non-word "Pal-Man", created by switching one letter (ק) with another letter (ל), following[50]. The displays' background was black, the letters were white, and were written in David font. The words (10.9° x 14.3) appeared in the middle of the screen.

In Experiment 5, three English words were used as primes: 'Triangle' served as a congruent semantic prime, 'Square' as an incongruent semantic prime, and 'Tsianmle' as a non-word prime (made by switching the triangle letters (r, g) with the letters (s, m), to create a meaningless word). The displays were presented on a black background and the letters were white in David font. The words (10.9° x 14.3°) appeared in the middle of the screen. The Kanizsa display used in Experiment 1 was used here as the left-side triangle Kanizsa display. The same display flipped horizontally was used to create a second Kanizsa display, in which the Kanizsa triangle appeared on the right side.

Apparatus

Experiments 1–4 were presented on a HZ0010 HooZo 10.1" tablet screen, using an Android application (operating with Android version 8.1.0). The experimental code was written using Java and JavaScript. Experimental condition per participant was randomly assigned by the code. Participants' responses were given by a tap on the intended touchscreen buttons. During the experiments, the tablet was in an arm-length distance (i.e., ~50 cm) from the participant's face, held slightly lower than eye-level, with an estimated visual angle of 15.0° x 25.7°. These viewing parameters allow participants to see the entire screen, aiming to reach a standardized visual field.

Procedure

Experiments 1–4 were conducted outside the lab, in natural environments. Participants were standing, and held the experiment tablet themselves. Experiment 5 was carried out online. In all five experiments, after giving their informed consent by selecting the 'I give my consent' option, participants were given short written instructions, explaining that two pictures would be briefly presented, followed by a related question. The participants were asked to focus on the fixation symbol in the middle of the screen (a red cross) throughout the experiment. They were further instructed to look carefully at the following pictures and answer the question as accurately and as quickly as they could.

The experiment began when the participants pressed the "start the experiment" button. A 1200 ms 3-2-1 countdown appeared (each digit was presented for 400 ms), preparing participants for the upcoming presentation of the stimuli. This was followed by a black screen with the fixation in the middle, presented for 300 ms. Afterwards, one of the prime displays (semantic/non-semantic in Experiments 1–4, congruent/incongruent/non-word in Experiment 5) appeared for 1000 ms, followed by another black screen with a fixation in the middle for 100 ms. Then, one of the test displays (Kanizsa display/no-Kanizsa display in Experiments 1–4, Kanizsa on the right/left in Experiment 5) was presented for 800 ms. The duration of the test stimulus was selected following pretesting, aimed at making sure that it is consciously perceived yet briefly presented, to make the test task more challenging, and hence more sensitive. At 800 ms, 80% of the participants in the pretest reported seeing the Kanizsa shape.

Subsequently, in Experiments 1-4, an answer displays appeared until the participant tapped on one of the answer option-buttons to the question "What shape did you see?". There were three answer options: Triangle - represented by a triangle Kanizsa drawing, Square - represented by a square Kanizsa drawing, and an "I didn't see any shape" option. In Experiment 5, participants were asked "on which side did you see a shape?". The following options were given: "Left", "Right" and "I didn't see any shape". The options were ordered in one of two orders (either "Left", "Right", "I didn't see any shape", or the other way around), counterbalanced between participants. RT was measured from the time the question was presented until selecting one of the three buttons. After completing the trial, participants were asked to fill

their age (in digits) and to self-report their sex (selecting between three options: Female/Male/Other). Each participant completed one trial. The entire experiment took ~2 min.

In experiments 1–4, after the trial ended, participants were verbally debriefed and asked what was the first thing that came to mind when taking the experiment. In experiments 1, 3 and 4, all said they saw a video game. Then, the experimenter asked if they knew which game it was; all but two participants confirmed they recognized the Pac-Man game and they are familiar with the game. In experiment 2, all participants said they saw a gift or a package.

Analysis

Results were analyzed using Chi tests. Assumptions of Chi tests were checked and me; specifically, all expected frequencies exceeded 5. Multiple comparisons were corrected for using the Tree Benjamini-Hochberg (TreeBH) method[51]. TreeBH takes into account the dependency structure of all the analyses of the manuscript (see Supplementary Fig. 2 in Material), and corrects p-values and alpha levels per family of analyses in a hierarchical manner. All p-values were corrected using the Simes method, assigning p-values to higher levels of the analyses based on the corrected p-values of their lower-level descendent analyses, as this method provides higher power. Note that we did not apply this correction if it turned a non-significant result into a significant one (which did not happen in the current paper). All p-values reported in the paper are corrected p-values. The effect sizes of Chi tests are reported using φ and Cramer's V for tests of a 2×2 structure or larger, respectively. Confidence intervals of 95% (95% CI) are also reported per each such analysis. In cases in which statistics were precisely equal to zero or one we reported these numbers without decimal points, to reflect this equivalence (e.g., $p = 1$).

In addition to these Null Hypothesis Significance Testing (NHST) methods, we analyzed our results in a Bayesian manner as well using R statistical software (version 4.4.3)[52]. Bayes Factors (BF) were calculated using the BayesFactor package (version 0.9.12.4.7)[53]. We used the independent multinomial with a fixed rows method[54], with the default prior concentration value (a) of 1 for the alternative hypothesis. Sensitivity analysis showed that as prior concentration increases from 1 to 100, BF values that supported the existence of an effect with the default prior showed weaker support for the effect, while BF values that supported the null effect with the default prior showed stronger support for the null. But importantly, these priors affected only the strength of the evidence, not the overall conclusion. We adopted the convention that a BF less than 0.1 implies strong evidence for the lack of an effect (i.e., H_0 is at least 10 times more likely than H_1 given the data), a BF between 0.1 and 0.33 provides moderate evidence for the lack of an effect, a BF between 0.33 and 3 suggests insensitivity of the data (anecdotal evidence for the lack or presence of an effect, for $0.33 < BF < 1$ or $1 < BF < 3$, respectively), a BF between 3 and 10 denotes moderate evidence for the presence of an effect (i.e., H_1), a BF between 10 and 100 implies strong evidence, and a BF greater than 100 suggests extreme evidence for the presence of an effect[55]. Last, we derived Bayesian estimates of the effect sizes (φ and Cramer's V) using the effect size package[56]. These estimates are also reported, along with their 95% credible interval (95% CrI).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Results

Kanizsa illusion reduction via image primes

As our experiments consisted of a single trial per participant, we calculated the percentage of correct and wrong answers in each prime group. In experiments 1–4, a "correct answer" was an answer that matched the Kanizsa display type (i.e., for the Kanizsa display, reports of 'triangle' are considered correct and reports of "square" or "no shape" wrong, while for the no-Kanizsa display, 'no-shape' is considered correct and the others wrong). In experiment 5, where the Kanizsa triangle could appear either on

Exp.1: Can a semantic prime image reduce Kanizsa perception?

Exp. 2: Is it simply attention engagement?

Exp. 3: Can a semantic prime word reduce Kanizsa perception?

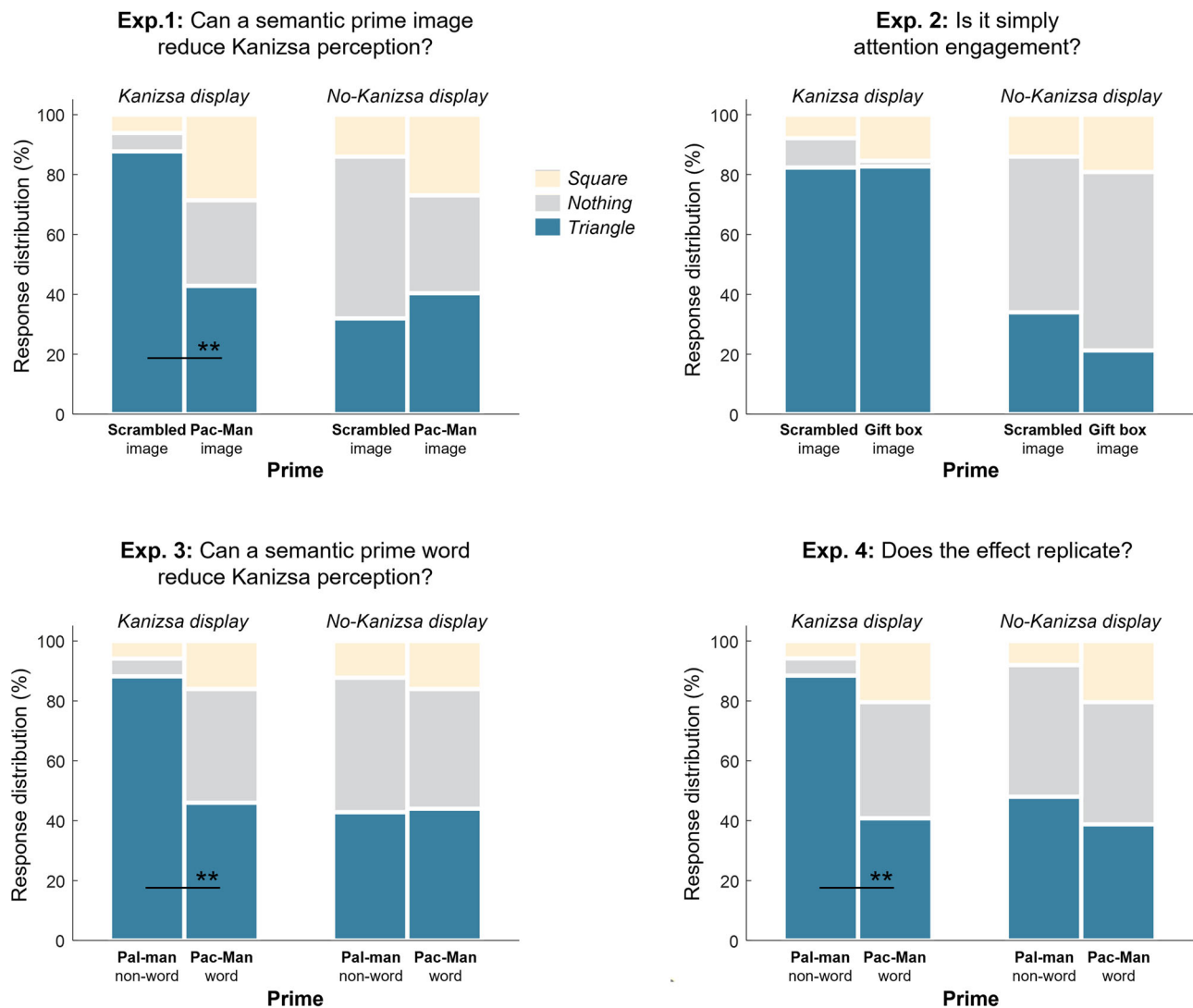**Exp. 4:** Does the effect replicate?

Fig. 2 | Results of experiments 1-4. For each experiment, the distribution of participants' responses is shown for the Kanizsa display on the left and the no-Kanizsa display on the right. Within each display type, the left column corresponds to the meaningless prime and the right column to the meaningful prime. High proportions of seeing a triangle (blue segments) occurred only when a Kanizsa triangle was displayed, and it was preceded by a meaningless or a non-related (Experiment 2) prime. When the prime was meaningful and related to Pac-Man, either as an image

(Experiment 1) or as a word (experiments 3 and 4), reports of seeing a triangle dropped to a comparable level as when no Kanizsa triangle was displayed. In Experiment 2, when the prime was meaningful but was not semantically relevant to Pac-Man (i.e., a gift box), this drop did not occur, suggesting it is indeed the semantic content of the prime that affected perception of the Kanizsa illusion. Each experiment included $n = 200$ participants. **$p < 0.001$.

the left side or the right side, a "correct answer" is one that matches the location in which the Kanizsa display appeared. Reaction times (RT) were also analyzed, yet with no prior hypothesis as this was not a speeded task. Indeed, no substantial effects were found in RT, and we focus here on the analyses of correct responses (but see Supplementary Note 1).

In Experiment 1, when a triangle Kanizsa display was presented, participants who received a Pac-Man gameboard prime were less than half as likely to report seeing a triangle than participants who were primed with a scrambled version of the same gameboard (42.9% vs. 87.8% respectively; $\chi^2(1, N = 98) = 21.80$; $p < 0.001$, $\varphi = 0.472$, 95% CI = [0.274, 0.670]; Fig. 2). These findings were supported by Bayesian analysis, providing extreme evidence for the presence of an effect[55] ($BF_{10} = 18248.884$, $\varphi = 0.454$, 95% CrI = [0.273, 0.609]). In the control condition, in which a no-Kanizsa display was presented, there was no difference between the groups in accuracy (i.e., reporting no shape; Pac-Man prime: 32.7%, scrambled prime: 54%; $\chi^2(1, N = 102) = 4.72$; $p = 0.654$, $\varphi = 0.215$, 95% CI = [0, 0.409], though with an inconclusive $BF_{10} = 2.473$, $\varphi = 0.209$, 95% CrI = [0.029, 0.382]). A post-hoc analysis showed that the semantic prime actually reduced seeing the

Kanizsa to the point that the ratio of reporting seeing nothing was not statistically different than that ratio when no Kanizsa shape was even displayed (reporting seeing nothing in the Kanizsa display with the Pac-Man prime vs. in the no-Kanizsa display with the Pac-Man prime: 28.6% vs. 32.7%; $\chi^2(1, N = 101) = 0.20$; $p = 0.654$, $\varphi = 0.045$, 95% CI = [0, 0.236]). There was moderate evidence that the two ratios were equivalent (i.e., that there was no difference between them $BF_{10} = 0.249$, $\varphi = 0.070$, 95% CrI = [0.003, 0.226]), implying that the semantic prime may have abolished the illusion altogether.

Yet an alternative interpretation is that participants failed to see the shape not due to the semantic activation evoked by the prime, but due to attentional engagement by the prime image[57]. Possibly, as the prime gameboard was more interesting than its scrambled version, it might have engaged participant's attention, leaving too few resources to detect the upcoming Kanizsa shape within the rich, crowded display. To test this interpretation, in Experiment 2 the prime was either an unrelated image (a gift box), created to resemble the gameboard used in Experiment 1 in terms of color and usage of dots, or a scrambled version of that image. Here,

participants were as likely to provide an accurate answer when the triangle Kanizsa display was presented (gift box prime: 82.7%; scrambled prime: 82.4%; $\chi^2(1, N = 103) < 0.01$; $p = 0.964$, $\varphi = 0.004$, 95% CI = [0, 0.006], $BF_{10} = 0.185$, $\varphi = 0.065$, 95% CrI = [0.003, 0.210]) and when it was not (gift box prime: 59.6%; scrambled prime: 52.0%; $\chi^2(1, N = 97) = 0.56$; $p = 0.453$, $\varphi = 0.076$, 95% CI = [0, 0.275], $BF_{10} = 0.326$, $\varphi = 0.088$, 95% CrI = [0.004, 0.265]; Fig. 2). Post-hoc analysis including both Experiment 1 and Experiment 2 showed that the abovementioned effects were manifested in a triple interaction between prime, display type and experiment in predicting reports of "triangle" ($F(1, 392) = 14.44$, $p < 0.001$, $\eta^2_p = 0.036$, 95% CI = [0.012, 1]). Taken together, these results suggest that the prime affected detecting the Kanizsa illusion in Experiment 1 and not in Experiment 2. Therefore, the effects found in Experiment 1 could not be easily explained by the mere meaningfulness of the prime, but rather suggest that they depend on the meaning of that prime being related to the Pac-Man game.

Kanizsa illusion reduction via word primes

Next, we tested if the effect can be evoked also by a verbal prime, to exclude any possibility that priming was driven by some perceptual features. In Experiment 3, the prime was accordingly the Hebrew word "פקמן" (Pac-Man), or a corresponding non-word "פלמן" (Pal-Man). Again, an effect was found, such that when a Kanizsa triangle display was presented, participants in the Pac-Man word prime group were less likely to report seeing a triangle than in the non-word prime group (46.0% vs. 88.2% respectively; $\chi^2(1, N = 101) = 20.47$; $p < 0.001$, $\varphi = 0.450$, 95% CI = [0.255, 0.645], $BF_{10} = 8651.865$, $\varphi = 0.435$, 95% CrI = [0.259, 0.588]; Fig. 2). Again, no difference in accuracy was found in the control condition, which contained no Kanizsa illusion (Pac-Man word prime: 40.0%, non-word prime: 44.9%; $\chi^2(1, N = 99) = 0.24$; $p = 0.622$, $\varphi = 0.050$, 95% CI = [0, 0.244], $BF_{10} = 0.274$, $\varphi = 0.073$, 95% CrI = [0.004, 0.238]).

Experiment 4 was a direct replication of Experiment 3. We chose to replicate this experiment since its results were most far-reaching with respect to the effect of cognition on perception: an effect of a word prime – that cannot be ascribed to any perceptual effect evoked by a prime image – on the likelihood of perceiving a visual illusion. Accordingly, we wanted to verify its reproducibility in another experiment, conducted on a new sample. Indeed, Experiment 4 fully replicated the results: here too, under the Kanizsa display, only 40.8% of participants reported a triangle following the Pac-Man word prime, compared to 88.5% following the non-word prime ($\chi^2(1, N = 101) = 25.29$; $p < 0.001$, $\varphi = 0.500$, 95% CI = [0.305, 0.695], $BF_{10} = 112043.032$, $\varphi = 0.484$, 95% CrI = [0.305, 0.632]), while no difference was found under the no-Kanizsa display (Pac-Man word prime: 40.8%, non-word prime: 44%; $\chi^2(1, N = 99) = 0.10$; $p = 0.749$, $\varphi = 0.032$, 95% CI = [0, 0.220], $BF_{10} = 0.256$, $\varphi = 0.067$, 95% CrI = [0.003, 0.223]; Fig. 2).

Kanizsa illusion enhancement

Finally, in Experiment 5 we set to examine if the effect can be reversed, so to enhance the tendency to see the Kanizsa shape following a congruent word prime indexing the expected shape ('Triangle'), as compared to an incongruent one indexing a different shape ('Square'), or a non-word. To make sure we are indeed probing Kanizsa perception and not simple response priming, here the task was not to report which shape, if at all, was present, but to detect if it appeared to the right or to the left side of the fixation cross (note that the same concern did not apply in Experiments 1-4, as there, the prime stimuli are not expected to bias participants towards choosing the "triangle" response, while here the prime itself is the word "triangle"). We reasoned that if priming increased illusory perception, participants will be more likely to correctly detect the shape location. This was indeed the case ($\chi^2(2, N = 348) = 18.28$; $p < 0.001$, Cramer's V = 0.229, 95% CI = [0.115, 0.329], $BF_{10} = 167.035$, Cramer's V = 0.232, 95% CrI = [0.128, 0.329]; Fig. 3). Post-hoc analyses showed that participants had higher accuracy in the congruent group compared with the incongruent group (50.0% vs. 24.1%, respectively; $\chi^2(1, N = 232) = 16.63$; $p < 0.001$, $\varphi = 0.268$, 95% CI = [0.139, 0.396],

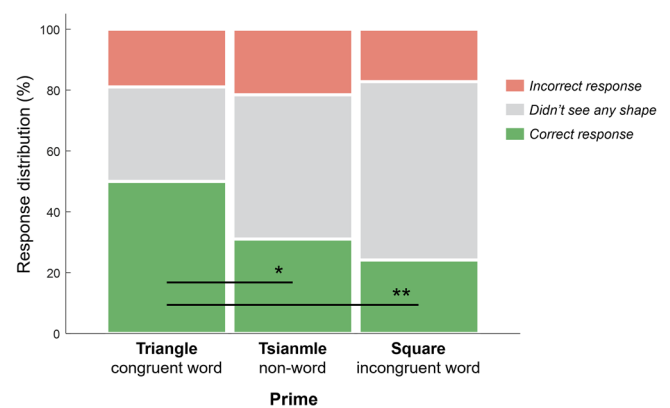**Exp. 5:** Can a semantic prime enhance Kanizsa perception?

Fig. 3 | Results of Experiment 5. Accurately detecting where the Kanizsa triangle appeared (green segments) was modulated by the prime group. The word "Triangle" as prime yielded higher accuracy than the non-word "Tsianmle" and the word "Square", suggesting that a semantic prime relating to the illusory shape enhanced its perception. $n = 348$ participants, *$p < 0.01$, **$p < 0.001$.

$BF_{10} = 674.715$, $\varphi = 0.263$, 95% CrI = [0.143, 0.383]), and also compared with the non-word group (31%; $\chi^2(1, N = 232) = 8.66$; $p = 0.005$, $\varphi = 0.193$, 95% CI = [0.064, 0.322], $BF_{10} = 12.021$, $\varphi = 0.190$, 95% CrI = [0.065, 0.311]). Accuracy was not found to differ between the non-word group and the incongruent group ($\chi^2(1, N = 232) = 1.38$; $p = 0.240$, $\varphi = 0.077$, 95% CI = [0, 0.206], $BF_{10} = 0.289$, $\varphi = 0.075$, 95% CrI = [0.004, 0.198]). It was also not found to be modulated by the side of the Kanizsa display ($\chi^2(1, N = 348) = 0$, $p = 1$, $\varphi = 0$, 95% CI = [0, 0], $BF_{10} = 0.127$, $\varphi = 0.037$, 95% CrI = [0.002, 0.119]).

Discussion

In this work, we show that participants can be semantically primed against seeing the Kanizsa shape (Experiments 1, 3 and 4), that this result cannot be explained by attentional engagement of a meaningful prime (Experiment 2), and that semantic priming can even enhance the probability of perceiving a Kanizsa shape (Experiment 5). Taken together, the results of the five experiments we conducted reveal a surprisingly strong phenomenon whereby the likelihood of seeing the Kanizsa illusion is either reduced or enhanced by semantic priming.

Previous work has demonstrated a modulatory effect of priming[41,42], of task[38] and of attention[39] on illusory perception, as well as affective influences on the probability of seeing an illusion[44]. Here, we demonstrate an abolishment of an illusion due to semantic priming. This suggests that visual illusions are not as impervious to cognition as previously thought[29,30], challenging claims of cognitive impenetrability[17–19] and supporting the role of top-down knowledge-based processes in perception[8,49,58–60], in line with predictive processing accounts[6,61].

According to these accounts, top-down signals modulate the activity in lower-level areas, allowing expectations and previous knowledge to affect processing[48]. Indeed, such top-down effects were reported at the earliest stages of the cortical visual hierarchy[62], specifically in the context of illusory contours. For example, single cell recordings in monkeys revealed that illusory contours evoke responses in V1 and V2 neurons[63,64], with the latter responding earlier to the contours than the former[65], suggesting the response reflects feedback connections. In humans, this effect was shown to specifically target deep layers of V1[66]. Later studies focused more generally on the influence of expectations on neural activity (for review, see ref. [67]). These studies further demonstrated that expectations about the probability of a stimulus to appear (e.g., when a cue predicts the appearance of a specific stimulus) bias neural activity prior to the presentation of that stimulus in various brain areas, making the stimulus easier to detect[68], and evoking

templates of the expected stimulus already in V1[69]. Arguably then, in our study, the presentation of the Pac-Man related stimulus might have evoked associated templates (i.e., of the Pac-Man figures), biasing participants' interpretations of the upcoming information towards perceiving them as Pac-Man figures rather than as background of the Kanizsa shape. This is in line with matching models of perception[48], which were further strengthened by EEG studies showing relatively early contextual effects of scene gist on object processing[49,70]. At this point, this is merely speculative; further research is needed to determine the underlying neural mechanisms of the effects we observed, especially given ongoing debates about the source of the effect of expectations (e.g., if they affect perceptual processing itself, or the decision criterion[71,72]), and how early in the visual hierarchy they occur – as our results, being behavioral, cannot speak to that question.

Can the current findings be criticized on the same methodological grounds directed at previous demonstrations of top-down effects[17,20]? In our case, we tested both confirmatory and disconformity predictions, showing that the effect also disappears when an irrelevant prime is used (Experiment 2). Second, to minimize the role of post-perceptual judgments, the stimuli were briefly presented, and the task (choosing a shape) directly examined perception (there was an objectively 'correct' answer to choose from three alternatives). Third, to avoid demand and response biases, our manipulation was not apparent, and each participant completed only one trial, thereby not knowing in advance what the question would be. Fourth, special care was taken to control for potential low-level confounds, both when designing the control conditions and the irrelevant prime in Experiment 2. Fifth, memory effects were minimized by immediately probing participants' perception after the display has been presented. Lastly, we showed that our effect cannot be simply explained by attentional engagement that is irrelevant to the semantic prime. Notably though, we concede that it might be mediated by some form of feature-based attention[73], guiding participants to perceive the local inducers, and not the global Kanizsa shape, as the figure rather than the ground[14,74,75]. Yet from our perspective, even if this is the case, this would not be a confound but rather part of the mechanism by which cognition affected perception[14,76], which is in line with similar views in the field[1,29,77].

Limitations

Naturally, the current study is not free of limitations and outstanding issues. First, the Kanizsa detection percentage in experiment 5 is lower than those of experiment 1–4, despite sharing similar stimuli. This might stem from the different task, or from the fact that Experiment 5 was conducted online rather than in person, which was the case in Experiments 1–4. Second, all our experiments were designed to yield suboptimal performance, via short presentation durations and crowding of the Kanizsa illusion that was placed peripherally rather than foveally. This was aimed at assuring variability in the dependent measure so that it would be sensitive to the manipulation (i.e., to avoid a ceiling effect). Our approach further accords with the claim that such impoverished presentation conditions might be more sensitive for finding effects of cognitive penetrability, as top-down effects are arguably more pronounced when the bottom-up stimulation is noisier[29]. Future work could systematically manipulate the noise level of the stimulus, testing the limits of cognitive penetrability as the stimulus grows closer to the typical Kanizsa display, presented in isolation and foveally. Finally, a puzzling result pertains to the tendency of a relatively large proportion of participants to report seeing a shape (triangle/square) even for the no-Kanizsa display. This response might be driven by demand characteristics[78]—that is, participants' assumption that they are expected to see something (as the question "which shape did you see?" already contains a bias), leading them to guess that one of the shape options is more feasible than the option of no shape at all. It is also possible that the no-Kanizsa display we used gave rise to some unplanned shapes, despite our efforts to design it otherwise. This echoes the well-established tendency to see shapes and patterns, even in meaningless stimuli[79]. Thus, the field will benefit from additional studies using different stimuli to further investigate the possible susceptibility of the Kanizsa illusion to top-down effects of cognition.

Data availability
The complete response data of all five experiments are publicly available at: https://osf.io/wmak2/.

Code availability
Analysis and experimental code used for this study are publicly available at: https://osf.io/wmak2/[80].

References
1. Lupyan, G. Cognitive penetrability of perception in the age of prediction: predictive systems are penetrable systems. *Rev. Philos. Psychol.* **6**, 547–569 (2015).
2. Stokes, D. Cognitive penetrability of perception. *Philos. Compass* **8**, 646–663 (2013).
3. Vetter, P. & Newen, A. Varieties of cognitive penetration in visual perception. *Conscious Cogn.* **27**, 62–75 (2014).
4. Dennett, D. C. Cognitive science as reverse engineering several meanings of "Top-down" and "Bottom-up". *Stud. Log. Found. Math.* **134**, 679–689 (1995).
5. Gilbert, C. D. & Sigman, M. Brain states: top-down influences in sensory processing. *Neuron* **54**, 677–696 (2007).
6. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 181–204 (2013).
7. O'Callaghan, C., Kveraga, K., Shine, J. M., Adams, R. B. & Bar, M. Predictions penetrate perception: converging insights from brain, behaviour and disorder. *Conscious Cogn.* **47**, 63–74 (2017).
8. Bar, M. The proactive brain: using analogies and associations to generate predictions. *Trends Cogn. Sci.* **11**, 280–9 (2007).
9. de Lange, F. P., Heilbron, M. & Kok, P. How do expectations shape perception? *Trends Cogn. Sci.* **22**, 764–779 (2018).
10. Churchland, P. M. Perceptual plasticity and theoretical neutrality: a reply to Jerry Fodor. *Philos. Sci.* **55**, 167–187 (1988).
11. Teufel, C., Dakin, S. C. & Fletcher, P. C. Prior object-knowledge sharpens properties of early visual feature-detectors. *Sci. Rep.* **8**, 1–12 (2018).
12. Kanai, R., Komura, Y., Shipp, S. & Friston, K. Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philos. Trans. R. Soc. B: Biol. Sci.* **370**, https://doi.org/10.1098/RSTB.2014.0169 (2015).
13. Seth, A. *Being You: A New Science of Consciousness* (Faber & Faber, 2021).
14. Stokes, D. Attention and the cognitive penetrability of perception. *Australas. J. Philos.* **96**, 303–318 (2018).
15. Brandman, T. & Peelen, M. V. Interaction between scene and object processing revealed by human fMRI and MEG decoding. *J. Neurosci.* **37**, 7700–7710 (2017).
16. Biderman, D., Shir, Y. & Mudrik, L. B or 13? Unconscious top-down contextual effects at the categorical but not the lexical level. *Psychol. Sci.* **31**, 663–677 (2020).
17. Firestone, C. & Scholl, B. J. Cognition does not affect perception: evaluating the evidence for "top-down" effects. *Behav. Brain Sci.* **39**, https://doi.org/10.1017/S0140525X15000965. (2016).
18. Fodor, J. A. *Modularity of Mind* (MIT Press, 1983).
19. Pylyshyn, Z. Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behav. Brain Sci.* **22**, 341–423 (1999).
20. Firestone, C. & Scholl, B. J. Can you experience 'top-down' effects on perception? The case of race categories and perceived lightness. *Psychon. Bull. Rev.* **22**, 694–700 (2015).

Article

21. Bhalla, M. & Proffitt, D. R. Visual–motor recalibration in geographical slant perception. *J. Exp. Psychol. Hum. Percept. Perform.* **25**, 1076–1096 (1999).

22. Durgin, F. H. et al. Who is being deceived? The experimental demands of wearing a backpack. *Psychon. Bull. Rev.* **16**, 964–969 (2009).

23. Levin, D. T. & Banaji, M. R. Distortions in the perceived lightness of faces: the role of race categories. *J. Exp. Psychol. Gen.* **135**, 501–512 (2006).

24. Levin, D. T., Baker, L. J. & Banaji, M. R. Cognition can affect perception: restating the evidence of a top-down effect. *Behav. Brain Sci.* **39**, e250 (2016).

25. Rolfs, M. & Dambacher, M. What draws the line between perception and cognition? *Behav. Brain Sci.* **39**, e257 (2016).

26. Vinson, D. W. et al. Perception, as you make it. *Behav. Brain Sci.* **39**, e260 (2016).

27. Lamme, V. A. F. Independent neural definitions of visual awareness and attention. in *Cognitive Penetrability of Perception: Attention, Action, Strategies, and Bottom-Up Constraints* 171–191 (Nova Science, 2005).

28. Adelson, E. H. Lightness perception and lightness illusions. in *The New Cognitive Neurosciences* 2nd edn (ed. Gazzaniga, M.) (MIT Press, 2000).

29. Lammers, N. A., de Haan, E. H. & Pinto, Y. No evidence of narrowly defined cognitive penetrability in unambiguous vision. *Front Psychol.* **8**, 232856 (2017).

30. van Buren, B. & Scholl, B. J. Visual illusions as a tool for dissociating seeing from thinking: a reply to Braddick (2018). *Perception* **47**, 999–1001 (2018).

31. Newen, A. & Vetter, P. Why cognitive penetration of our perceptual experience is still the most plausible account. *Conscious Cogn.* **47**, 26–37 (2017).

32. Müller-Lyer, F. C. Optische urteilstäuschungen. *Arch. Anat. Physiol. Physiol. Abt.* **2**, 263–270 (1889).

33. Rivers, W. H. R. Observations on the senses of the Todas. *Br. J. Psychol.* **1**, 321–396 (1905).

34. Ebbinghaus, H. *Grundzüge Der Psychologie* (Verlag von Viet & Co, 1902).

35. de Fockert, J., Davidoff, J., Fagot, J., Parron, C. & Goldstein, J. More accurate size contrast judgments in the ebbinghaus illusion by a remote culture. *J. Exp. Psychol. Hum. Percept. Perform.* **33**, 738–742 (2007).

36. Bremner, A. J. et al. Effects of culture and the urban environment on the development of the Ebbinghaus illusion. *Child Dev.* **87**, 962–981 (2016).

37. Caparos, S. & Boissin, E. The relationships between urbanicity, general cognitive ability, and susceptibility to the Ebbinghaus illusion. *Psychol. Res* **88**, 1540–1549 (2024).

38. Weidner, R. & Fink, G. R. The neural mechanisms underlying the Müller-Lyer illusion and its interaction with visuospatial judgments. *Cereb. Cortex* **17**, 878–884 (2007).

39. Tsal, Y. A Mueller–Lyer illusion induced by selective attention. *Q. J. Exp. Psychol. Sect. A* **36**, 319–333 (1984).

40. Ternus, J. The problem of phenomenal identity. *A Source Book of Gestalt Psychology*. 149–160 (Routledge, 2007).

41. Yu, K. Can semantic knowledge influence motion correspondence? *Perception* **29**, 693–707 (2000).

42. Hsu, P., Taylor, J. E. T. & Pratt, J. Frogs jump forward: semantic knowledge influences the perception of element motion in the ternus display. *Perception* **44**, 779–789 (2015).

43. Otten, M., Pinto, Y., Paffen, C. L. E., Seth, A. K. & Kanai, R. The uniformity illusion. *Psychol. Sci.* **28**, 56–68 (2017).

44. Kraus, N., Niedeggen, M. & Hesselmann, G. Negative affect impedes perceptual filling-in in the uniformity illusion. *Conscious. Cogn.* **98**, 103258 (2022)

45. Kanizsa, G. Margini quasi-percettivi in campi con stimolazione omogenea. *Riv. Psicol. (1912)* **49**, 7–30 (1955).

46. Rubin, E. *Visuell Wahrgenommene Figuren* (Glydenalske, 1921).

47. Rock, I. & Anson, R. Illusory contours as the solution to a problem. *Perception* **8**, 665–681 (1979).

48. Bar, M. Visual objects in context. *Nat. Rev. Neurosci.* **5**, 617–29 (2004).

49. Truman, A. & Mudrik, L. Are incongruent objects harder to identify? The functional significance of the N300 component. *Neuropsychologia* **117**, 222–232 (2018).

50. Frost, R. & Yogev, O. Orthographic and phonological computation in visual word recognition: evidence from backward masking in Hebrew. *Psychon. Bull. Rev.* **8**, 524–530 (2001).

51. Bogomolov, M., Peterson, C. B., Benjamini, Y. & Sabatti, C. Hypotheses on a tree: new error rates and testing strategies. *Biometrika* **108**, 575–590 (2021).

52. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2021).

53. Morey, R. D. & Rouder, J. N. BayesFactor: computation of bayes factors for common designs. *CRAN: Contributed Packages* Morey. Package 'bayesfactor' (2015). https://doi.org/10.32614/CRAN.package.BayesFactor.

54. Gunel, E. & Dickey, J. Bayes factors for independence in contingency tables. *Biometrika* **61**, 545–557 (1974).

55. Lee, M. D. & Wagenmakers, E.-J. *Bayesian Cognitive Modeling: A Practical Course* https://doi.org/10.1017/CBO9781139087759 (Cambridge University Press, 2014).

56. Ben-Shachar, M., Lüdecke, D. & Makowski, D. effectsize: estimation of effect size indices and standardized parameters. *J. Open Source Softw.* **5**, 2815 (2020).

57. Mack, A., Pappas, Z., Silverman, M. & Gay, R. What we see: Inattention and the capture of attention by meaning. *Conscious Cogn.* **11**, 488–506 (2002).

58. Hansen, T., Olkkonen, M., Walter, S. & Gegenfurtner, K. R. Memory modulates color appearance. *Nat. Neurosci.* **9**, 1367–1368 (2006).

59. Neri, P. Semantic control of feature extraction from natural scenes. *J. Neurosci.* **34**, 2374–2388 (2014).

60. Brooks, J. A. & Freeman, J. B. Conceptual knowledge predicts the representational structure of facial emotion perception. *Nat. Hum. Behav.* **2**, 581–591 (2018).

61. Hohwy, J. New directions in predictive processing. *Mind Lang.* **35**, 209–223 (2020).

62. Lee, T. S. & Mumford, D. Hierarchical Bayesian inference in the visual cortex. *J. Optical Soc. Am. A* **20**, 1434–1448 (2003).

63. Grosof, D. H., Shapley, R. M. & Hawken, M. J. Macaque VI neurons can signal 'illusory' contours. *Nature* **365**, 550–552 (1993).

64. von der Heydt, R., Peterhans, E. & Baumgartner, G. Illusory contours and cortical neuron responses. *Science (1979)* **224**, 1260–1262 (1984).

65. Lee, T. S. & Nguyen, M. Dynamics of subjective contour formation in the early visual cortex. *Proc. Natl. Acad. Sci. USA* **98**, 1907–1911 (2001).

66. Kok, P., Bains, L. J., van Mourik, T., Norris, D. G. & de Lange, F. P. Selective activation of the deep layers of the human primary visual cortex by top-down feedback. *Curr. Biol.* **26**, 371–376 (2016).

67. Summerfield, C. & de Lange, F. P. Expectation in perceptual decision making: neural and computational mechanisms. *Nat. Rev. Neurosci.* **15**, 745–756 (2014).

68. de Lange, F. P., Rahnev, D. A., Donner, T. H. & Lau, H. Prestimulus oscillatory activity over motor cortex reflects perceptual expectations. *J. Neurosci.* **33**, 1400–1410 (2013).

69. Kok, P., Failing, M. F. & de Lange, F. P. Prior expectations evoke stimulus templates in the primary visual cortex. *J. Cogn. Neurosci.* **26**, 1546–1554 (2014).

70. Mudrik, L., Lamy, D. & Deouell, L. Y. ERP evidence for context congruity effects during simultaneous object–scene processing. *Neuropsychologia* **48**, 507–517 (2010).

71. Bang, J. W. & Rahnev, D. Stimulus expectation alters decision criterion but not sensory signal in perceptual decision making. *Sci. Rep.* **7**, 17072 (2017).

72. Rungratsameetaweemana, N., Itthipuripat, S., Salazar, A. & Serences, J. T. Expectations do not alter early sensory processing during perceptual decision-making. *J. Neurosci.* **38**, 5632–5648 (2018).

73. Maunsell, J. H. R. & Treue, S. Feature-based attention in visual cortex. *Trends Neurosci.* **29**, 317–322 (2006).

74. Weissman, D. H., Mangun, G. R. & Woldorff, M. G. A role for top-down attentional orienting during interference between global and local aspects of hierarchical stimuli. *Neuroimage* **17**, 1266–1276 (2002).

75. Aydin, M., Herzog, M. H. & Öğmen, H. Attention modulates spatio-temporal grouping. *Vis. Res* **51**, 435–446 (2011).

76. Marchi, F. Attention and cognitive penetrability: the epistemic consequences of attention as a form of metacognitive regulation. *Conscious Cogn.* **47**, 48–62 (2017).

77. Lupyan, G. Changing what you see by changing what you know: the role of attention. *Front Psychol.* **8**, 250363 (2017).

78. Orne, M. T. Demand characteristics and the concept of quasi-controls. in *Artifacts in Behavioral Research* (eds. Rosenthal, R. & Rosnow, R. L.) 110–137 (Oxford University Press, 2009).

79. Voss, J. L., Federmeier, K. D. & Paller, K. A. The potato chip really does look like Elvis! Neural hallmarks of conceptual processing associated with finding novel shapes subjectively meaningful. *Cereb. Cortex* **22**, 2354–2364 (2012).

80. Davidson Litvak, N., Tal, A. & Mudrik, L. Cognition, Pac-Man and Perception: The Effect of Context on the Kanizsa Illusion. https://doi.org/10.17605/osf.io/wmak2 (2019)

Author contributions
N.D.L.: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing - original draft, Writing—review & editing; A.T.: Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing—original draft, Writing—review & editing; L.M.: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing - original draft, Writing—review & editing.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44271-025-00268-9.

Correspondence and requests for materials should be addressed to Liad Mudrik.

Peer review information *Communications Psychology* thanks Su-Ling Yeh, Gary Lupyan, Rasha Abdel Rahman, and Chien-Chun Yang for their contribution to the peer review of this work. Primary Handling Editors: Troby Ka-Yan Lui. [A peer review file is available].

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.