Donders Graduate School for Cognitive Neuroscience
Master of Science Programme
MSc Thesis

How Our Predictions do not Deceive Us:
An Investigation of the Illusory Perception of Upside-down Letters

by

Jochem Koopmans

Supervised by:

1. prof. dr. Marius V. Peelen
2. Sushrut Thorat, M.Sc.
3. Genevieve L. Y. J. Quek

Radboud University Nijmegen

Date of final oral examination: 8 July 2022

Student name: Jochem Koopmans

Student number: s4529901

Contents

Abstract: Predictive processing (PP) theories explain how the brain processes visual information. They have been used to explain why people presented with a display of letters have the experience of seeing more letters than they can report. The partial awareness hypothesis (PAH) suggests that predictions about the content enhance relatively poor visual representations letters in such displays, such that the richness is inferred rather than really experienced. Critical support comes from a study purportedly showing that upside-down letters can be illusorily perceived as upright under difficult viewing conditions (de Gardelle et al., 2009). Here, three experiments investigate whether this finding could be due to a response bias instead of illusory perception. In the first experiment, a replication of the original study showing illusory perception shows that upside-down letters are reported as upright as much as not-shown letters. However, this could be explained by only a few letters being represented in iconic memory (IM). The second experiment repeats this finding when more letters are maintained in IM, providing stronger evidence against the illusion effect. The third experiment, a full-power replication of the second, confirms that upside-down letters are not reported as upright any more than not-shown control letters. This suggests that the tendency to report letters as upright is a post-perceptual effect rather than a perceptual illusion, but can still be explained within the PP framework. The rejection of the illusory perception phenomenon revives the debate about whether visual consciousness overflows capacities of report.

Keywords: iconic memory, illusion, partial awareness hypothesis, partial report, predictive processing

Introduction

Human visual perception is often modelled as hierarchical predictive processing (e.g. Friston, 2018; Walsh et al., 2020). Predictive processing (PP) treats the brain as a comparator that builds an internal model of the environment and compares it against sensory input. Such an environment-model consists of different layers of complexity, corresponding to the complexity of information that neurons in the visual processing areas respond to.

An early visual processing area like V1 represents information pertaining to orientations, luminance, and spatial frequencies of luminance (e.g. Benvenuti et al., 2018; El-Shamayleh et al., 2013; Spratling, 2010), whereas at a later stage like in the fusiform face area (FFA) such and other features are combined to represent more complex entities like faces (Kanwisher & Yovel, 2006) or frequently perceived objects (Burns et al., 2019). Although hierarchical PP emphasizes the sequential nature of visual processing in explaining brain functioning, it does not preclude the existence of parallel processing (Hegdé & Felleman, 2007).

PP accounts of visual perception hold that at every processing stage, top-down predictions and bottom-up signals are integrated (Clark, 2013; Walsh et al., 2020). Top-down predictions are the predicted activation of a lower-level area given what the higher-level area currently represents. For example, if the FFA's activity is such that it can be said to represent a face, it will 'predict' colour-processing areas to show activity corresponding to a perception of skin colours and 'predict' areas processing geometric shapes to show activity corresponding to curvilinear features.

Such top-down predictions are compared to bottom-up sensory signals (Summerfield & de Lange, 2014). If indeed one continues to see a face, the predicted activity of the colour- and geometric shape-processing areas will match the activity driven by sensory information. The prediction and sensory evidence can be integrated without resulting in an unexplained 'remainder' activation. If, unexpectedly, a different stimulus appears (e.g., a bus passes in front of the face), there will be different colours and more rectilinear shapes to process. The predicted and sensorily driven activation of that area will differ greatly. This difference, or error, is propagated upward, where it corrects future expectations of seeing a face to seeing more of the bus.

The precise way in which predictions (termed 'priors' in PP parlance) and sensory information are integrated differs between PP models, but a shared characteristic between them is that they use Bayesian principles (de Lange et al., 2018; Summerfield & de Lange, 2014). What a certain visual processing area represents at a given time is therefore the result of integrating predicted activation with sensory information. In addition, the integration of these sources is weighted, such that imprecise predictions have a relatively small influence on the resulting representation when sensory evidence is precise, and vice versa. Strong predictions can therefore disambiguate an unclear stimulus (Brandman & Peelen, 2017).

PP has been used to explain the apparent richness of visual experience in the face of limited processing resources. According to the partial awareness hypothesis, limited sensory information can be combined with strong priors to drive an illusion of rich sensory awareness (Kouider et al., 2010). For example, pseudowords that are flashed very quickly can lead to subjective feeling of having seen a real word (Kouider & Dupoux, 2004). Presumably, strong predictions regarding the stimuli's lexicality modulate the weak sensory input to arrive at an illusory percept of linguistic entities.

Another way to ostensibly dissociate visual experience from processing resources is the Sperling paradigm (e.g. Sligte et al., 2010), named after Sperlings (1960) seminal experiments. Sperling found that subjects could only freely recall up to three or four items (letters or numbers) from a briefly flashed grid containing up to twelve items. In the crucial experiment, Sperling presented a tone after grid offset to indicate which row subjects were to report. Because every row was equally probable to be cued, this partial report procedure reduced processing related to reporting items to the cued row's four items while still requiring subjects to attend to all the rows during grid display. Subjects were able to report almost all the items in the cued row, indicating that information about nearly all the grid's letters was available for report (e.g. Block, 2011).

A different interpretation of this finding is that subjects are merely partially aware of the items in the grid (partial awareness hypothesis; Kouider et al., 2010). Instead of perceiving every individual letter, degraded processing of the letters could be integrated with predictions of letters being upright, generating the experience of seeing a grid with twelve letters. Specific attentional processes, e.g. directed by a cue, would provide additional processing resources to individuate up to four or five letters.

This hypothesis was investigated by de Gardelle et al. (2009), who hid an upside-down letter in a row that was subsequently not cued. After subjects reported the letters in the cued row, they were asked to select items which they thought had occurred in the grid, from a list containing both upright and upside-down letters. Subjects reported selected the upright version of the upside-down letter that had been shown in the grid more than its veridical upside-down alternative, even when controlling for false-alarms. The authors interpret this as showing that the items in non-cued rows are not processed in detail. Following the partial awareness hypothesis, they conclude that top-down predictions 'rotate' the imprecise sensory information of the upside-down letter to be upright (Kouider et al., 2010).

An alternative interpretation of these results maintains that upside-down letters need not be illusorily perceived. Instead, results could be based on a general response bias that favours selecting upright letters over upside-down ones. There are two main reasons why this alternative view is plausible, which will be described below.

First, the upside-down letter may not be maintained in visual processing up to the letter selection part of a trial. The visual representation of items after grid offset is thought to be maintained in iconic memory (IM), which has a large capacity but is very sensitive to overwriting by new visual input (Becker et al., 2013; Öğmen & Herzog, 2016; Rensink, 2014). Importantly, the mask presented at grid offset in de Gardelle et al. (2009) would destroy the iconic memory of the grid (Pinto et al., 2013; Sligte et al., 2008; Sligte et al., 2010).

This idea is corroborated by the (partial) recall performance in de Gardelle et al. (2009), which shows that subjects only correctly recalled only 1.16-1.49 out of the 4 items in the cued row. Since subjects did not know beforehand which out of the 3 rows would be cued, these recalled items suggest that roughly (1.49 * 3 ≈) 4.5 items were available from the whole grid of 12. This availability corresponds to the free report performance reported in other studies, in which subjects were asked to report all twelve letters (Dick, 1969; Sperling, 1960). It is possible that the mask in de Gardelle et al. (2009), presented prior to the cue, destroyed items that were not attended during the grid display but that would have been represented in IM.

Even if subjects initially attended the upside-down letter, the cue would favour attention to letters that did occur in the cued row and inhibit attention to IM letter representations of letters from other rows. This makes it less likely that the letter representation of the upside-down letter, whether it is upright or upside-down at that point, can be recognised in the list of letter options at the letter selection part of the trial that occurred afterwards.

Secondly, predictive processing modulations of the iconic memory representation of an upside-down letter are more likely to lead to the selection of other upright letters than the upside-down one. Iconic memory has been related to decaying V1 activity (Teeuwen et al., 2021). Considering that V1 represents low-level visual features, predictions of stimuli being linguistic entities (i.e. upright letters) would translate to activation patterns reflecting luminance, spatial frequencies, and contrast: properties that are represented in V1 activation patterns (Benvenuti et al., 2018) rather than whole letters (Quilty-Dunn, 2020). Upright and upside-down letters do not differ in such low-level properties since they are composed of the very same line segments which are internally identically arranged. Top-down predictions would therefore not influence iconic memory in a way that would make the activation caused by single letters to rotate to an upright orientation. An upside-down letter would be more likely to be (incorrectly) ‘disambiguated’ as another visually similar letter than its upright orientation (e.g. an upside-down V shares more low-level visual features with an upright A than an upright V).

These considerations invite a further investigation of the claimed illusory perception of upside-down letters as upright. The current research will investigate to what extent a response bias could explain de Gardelle et al.'s (2009) findings. To do so, three partial replications of their procedure will compare the possible illusion effect of selecting the upside-down letter in its upright orientation to important controls.

In experiment 1, the tendency to select the upside-down letter in its upright orientation will be compared against an upright shown and a not-shown control letter. The mask used in de Gardelle et al. (2009) will be omitted to increase the influence of top-down predictions on the iconic memory representation of the upside-down letter. If the upside-down letter is indeed illusorily perceived, it should be distinguishable from the tendency to report a not-shown letter as upright, but if the upside-down letter's orientation is not maintained (veridically or otherwise) its proportion of being selected as upright should be comparable to that of a not-shown letter. Experiments 2 and 3 (a full-powered replication of Experiment 2) employ a less taxing letter selection procedure that should be more sensitive to illusory upright representations of upside-down letters. In addition, the grid presentation time is reduced to further increase the influence of top-down predictions on the upside-down letter representation.

Experiment 1

The first experiment is procedurally similar to de Gardelle et al.'s (2009) study. Subjects will be shown grids of items that include an upside-down letter, recall the letters in the cued row and indicate which items from a list they recognise from the grid. The main goal is to compare reports of the upside-down letter to control items in order to investigate whether an illusion effect could also be explained as a response bias that would apply equally to not-shown letters. The upright shown item is included to assess the relative strength of a possible illusionary perception of the upside-down letter.

A secondary goal of this experiment will be to assess the role of attending items. The mask used in de Gardelle et al. (2009) is likely to overwrite representations of unattended letters, but might have the additional function of preventing the upside-down letter from attracting attention as a pop-out item. Such attention is known to improve the spatial resolution of perception (Anton-Erxleben & Carrasco, 2013), probably reducing the likelihood of the upside-down letter being perceived as upright. To assess possible upside-down letter saliency effects, there will be a condition without a cue to direct attention (free recall). Without a cue to direct attention away from the upside-down letter, the upside-down letter will be attended more if it does attract attention. As a consequence, it will also likely be more accurately represented and recalled, going against its illusory perception as upright.

Methods

Participants

The study was hosted on Prolific, an online experiment database (https://prolific.co/) and restricted to European subjects using a single computer screen. Subjects signed an electronic consent form and were compensated with GB£4.44 for their participation. Forty-five subjects completed the online task. Of these, one did not provide serious answers and eight indicated to have ignored letter orientation during the recognition question (see stimuli and procedure). This left 36 subjects for the analysis (20 male, μ_{age}= 31.50, SD_{age}= 11.71).

Procedure

The experiment consisted of four blocks containing thirty-two trials each. Each trial had five steps (see also Figure 1): a fixation cross (500ms), a 3x4 grid with upright letters, an upside-down letter, and a number (300ms), a cue or blank screen (300ms), followed by the recall question (self-paced) and recognition question (self-paced). The recall question asked subjects to enter the items in the cued row if it was a cued recall block or all the items in the grid if it was a free recall block. Subjects were instructed to enter any upside-down letters using their upright counterparts on the keyboard. The order in which subjects entered letters was not taken into account. The recognition question showed subjects four items and asked subjects to select the one(s) that they thought had occurred in the preceding grid. Trials were separated by a 750ms intertrial interval. A trial's recall type was alternated in 32-trial blocks, the trial type of the first block being randomised for each participant.

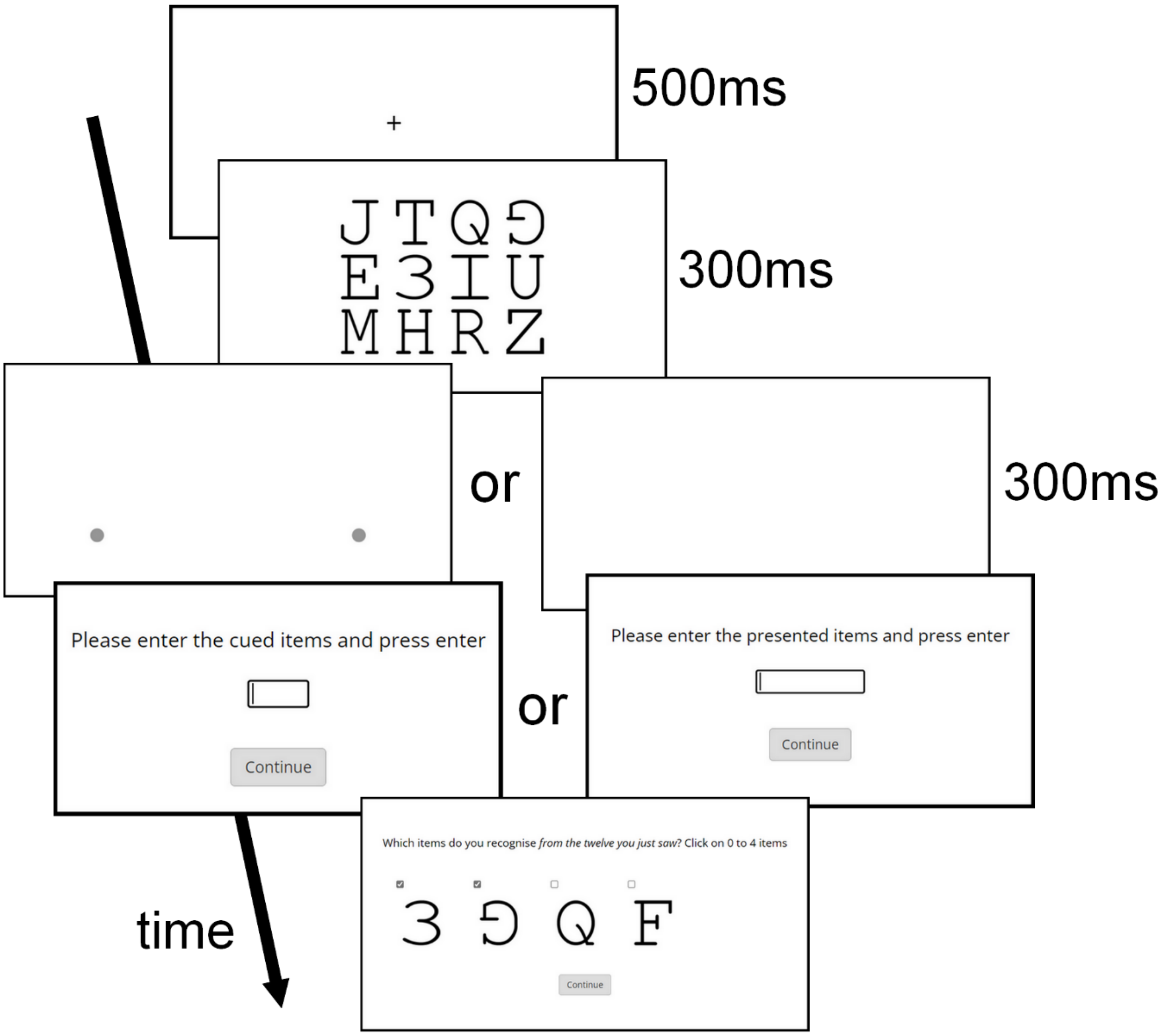
Each block started with an instruction about the upcoming trial type, followed by the 32 trials. The first two trials were considered practice and hence discarded prior to analysis. At the end of each block, subjects were given the opportunity to rest and they saw how well they performed on both the recall question (percentage of cued/presented items they entered correctly) and the recognition question (percentage of selected items that had indeed been shown in the grid).

Stimuli

All letters and numbers in the trials were of the black 'Courier New' monospaced font, letters were capitalised. This font was chosen for its readability and being monospaced. Upside-down letters were created by rotating an upright letter by 180 degrees. All letters and numbers were

Figure 1

An Example Trial in Experiment 1



Note. Trial structure for Experiment 1. Stimuli are not to scale. In the cued condition (left) subjects should enter 'MHRZ' (corresponding to the third row, cued by the dots). In the full report condition (right), all twelve items should be entered. In the final recognition task, the first three options should be selected.

edited to be 650x550 pixels and centralised in an 800x800pixel transparent-background (png) file. This letter pixel ratio corresponded best to most letters' height-width ratio. The letter Q was constructed manually by combining an O with the lower-right element of the letter K so as to maintain a similar height-width ratio as most other letters.

The 3x4 grid contained 9 upright letters, supplemented by a set of three ‘special’ items that served as input for the recognition question. Upright letters could be any of the English alphabet’s letters. The special items consisted of a single-digit number, an upside-down letter, and an upright letter. Letters in the special item set were selected from a subset of all the letters in the English alphabet which had unique visual characteristics when upside-down. An H was therefore not part of the subset, but a G was. The subset contained 15 letters.

In a block with cued recall trials, the grid was followed by a cue. The cue consisted of two grey dots that appeared to the left and right of a certain row and indicated the four items that needed to be remembered. The grey cue colour and placement outside the row were chosen to prevent backward masking or overwriting of the letters in that row (Black & Barbee, 1985). A visual cue instead of a more common auditory one was used to avoid further inter-subject variability in this online-administered task. Partial report superiority over free report has been found using visual cues (e.g. Black & Barbee, 1985; Sligte et al., 2008). The cued row was randomised while ensuring that each row was cued equally often per block. A trial in a free recall block showed an empty white screen instead of the cue.

The recognition question appeared only once the subject confirmed their answer on the recall question. It always showed four item probes, the subject’s task was to select which ones had been part of the preceding grid. These items were chosen to probe the subject’s memory about the grid’s three special items (number, upside-down-letter, upright-letter) and a letter that was not shown in the grid (not-shown-letter probe, shown in order in Figure 1). The number probe could either be correct or incorrect (i.e. a different number), the three letter probes could either be upright or upside-down, so each probe type had two alternatives. The alternative that was shown was randomised per probe on every trial, but counterbalanced such that both alternatives were shown on half of the trials. The probe order on screen was randomised on every trial.

Data-analysis

Data analysis was organised such that the background requirements for a possible illusion effect to occur were checked first. First, a comparison between recalled items on cued and free recall questions could indicate to what extent the grid’s letters were represented in IM. Subject entries to the recall question were counted as correct if they had indeed been in the cued row (cued recall) or in the grid (free recall). The number of correct entries was averaged per block, which, for free recall reflects item availability for the whole grid.

For cued recall, the number of correctly recalled items only shows availability of items for the cued part of the grid (1 out of 3 rows or 4 out of 12 letters). Since participants did not know ahead of time which row will be cued for recall, item availability for the whole grid can be reconstructed by correcting for the partial report size. The reconstructed availability was calculated by multiplying the recalled letters in the cued row by the number of rows (3). A superior reconstructed availability for cued recall over the availability of items in free recall is an indication that more items are retained in iconic memory than that can be reported (Sperling, 1960).

The recognition question asked subjects to indicate which of the probes they think was present in the grid and is of the main interest in this experiment. However, given the fragility of IM and the relatively long time between the grid display and the recognition question, it was worth investigating whether any information regarding the upside-down item is retained up to that point.

The first background check investigated how many recognition probes were selected on average. If the delay between the grid and the recognition question was too long, subjects might not remember much about the grid or be unconfident about their memory and therefore select close to no probes. In addition, this analysis allowed a first exploration of the saliency of special items when attention is directed away from them on cued recall trials compared to undirected attention in free recall trials. Such saliency should be reflected in an increase in the number of probe selections on free compared to cued recall trials.

The second background check investigated whether subjects specifically selected shown probe types, irrespective of their orientation, more than the not-shown probe type. To investigate this, selection rates for the correct number probe, and both alternatives for the upside-down-letter probe, upright-letter probe and not-shown-letter probe were calculated. If subjects forgot which items had been shown, selection rates for those probes would be comparable to that for the not-shown-letter probe. A comparison would therefore also reveal the extent to which the special items automatically grabbed attention between cued and free recall trials.

The third background analysis compared probe selection accuracy. Selection accuracy was determined for every probe type as the proportion of hits (correctly selected probe alternatives) and correct rejections (not selecting incorrect probe alternatives). The selection accuracy for the not-shown-letter probe can be derived from the selection rate calculated in the previous analysis, so it is not considered separately in this analysis. By not presenting a mask after the grid, it is possible that subjects accurately remember which items had been part of the grid. If the previous analyses indicated that special items attract attention, this analysis would reveal whether such attention leads to more accurate memory of them.

Finally, the last analysis calculated the proportion of upright probe alternatives among the times that each probe was selected, and compared this proportion between the three letter probes (upside-down-letter, upright-letter and not-shown-letter). The not-shown-letter probe's proportion of selecting the upright compared to upside-down alternative indicated the general tendency to report letters as upright, against which the same proportion for the upside-down-letter probe could be compared. The upright-letter probe's proportion of upright selections served to compare how strong a possible illusion would be when compared to an actually upright presented letter. The illusion effect was compared between cued and free recall trials, again because of a possible attentional effect on upside-down letter representations.

Results

Recall question

The recall question served to compare the availability of the items in the grid between cued and free recall trials and estimate how many items were represented in IM. A repeated-measures (RM) ANOVA on the (reconstructed) number of items available for report revealed no effect of recall type ($F(1,35) = 1.956$, $p = .171$), a marginal effect of block ($F(1,35) = 3.642$, $p = .065$, partial $\eta^2 = .094$) and no interaction effect ($F(1,35) = 2.404$, $p = .130$; see Figure 2). The (reconstructed) mean number of available items during recall did not differ between cued (4.8264, $SE_{mean} = 0.245$) and free recall trials (4.526, $SE_{mean} = 0.166$) and did not improve over time for either recall type.

The absence of an effect of recall type seemed due to a low performance on cued recall trials. The mean reconstructed availability of 4.8 out of 12 items (40.2%) was much lower than the 8-9 (65-75%) reported elsewhere (Sperling, 1960; see Figure 2). Cued recall performance was roughly equivalent to that reported in de Gardelle et al. (2009), who reported a reconstructed availability of up to 4.5 items. On free recall, subjects showed an availability of 4.5 out of 12 items (37.7%), which corresponds to the values found in other studies (Sperling, 1960).

Recognition question

Subjects selected on average 1.087 ($SE_{mean} = 0.082$) out of the four recognition probes per trial. The number of selected probes differed marginally between trial types ($F(1,35) = 3.101$, p=

.087, partial η^2= .081), with slightly more probes selected on free recall trials (1.102, SE_{mean}= 0.090) than on cued recall trials (1.071, SE_{mean}= 0.082). There was no effect of block (F(1,35)= 0.314, p= .579), nor an interaction between trial type and block (F(1,35)= 0.028, p= .868). Since there was no effect of block on the number of selections, block will not be included as a separate variable in the analyses that follow. In both conditions, subjects seemed confident enough to decide which probes had been shown.

The second control analysis tested whether subjects retained information regarding the items presented in the grid by comparing how often the different probe types were selected (in either orientation for the letter probes). If they did not, the not-shown-letter probe would be selected as often as the other probe types. This comparison was performed in a 4*2 RM ANOVA on selection proportion with factors probe type (correct number, upside-down-letter, normal-letter, not-shown letter) and recall type (cued, free recall). For the number probe, only selections for its correct alternative were included, since the incorrect number was not shown in the grid and could therefore not be remembered. The not-shown letter probe served as a control.

The analysis repeated the previously found medium-size[1] difference in selected items between cued and free recall trials (F(1,35)= 5.187, p= .029, partial η^2= .129), with more probes being selected more in free recall trials (each probe on average .259 times per trial, SE_{mean}= .016) than in cued recall trials (.233, SE_{mean}= .020). Importantly, there was a large effect of probe type (F(3,33)= 30.908, p< .001, partial η^2= .738), as shown in Figure 3a. The correct number probe (.202, SE_{mean}= .010) was selected as often as the not-shown-letter probe (.200, SE_{mean}= .022; p= .916). The upside-down- and upright-letter probes were also selected similarly often (respectively .298, SE_{mean}= .024 and .284, SE_{mean}= .021; p= .180). The upside-down- and upright-letter probes were both selected more than the correct number probe and the not-shown letter probe (all ps< .001).

A large significant interaction between trial type and probe (F(3,33)= 14.422, p< .001, partial η^2= .567) was driven by a reversal of the general effect probes being selected more often on the free recall trials for the not-shown letter probe. Where the other probes were selected more on free recall trials than on cued recall trials (ps< .020), the not-shown letter probe was selected more on cued compared to free recall trials (p= .002).

The third analysis compared the accuracy of decisions regarding the probe types, which would reveal whether subjects had remembered the special items in detail (see Figure 3b). Here, we

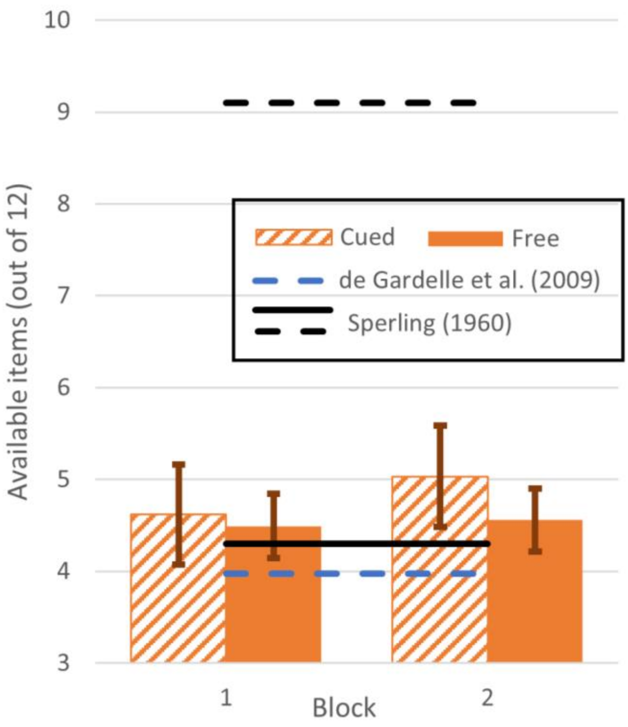
[1] Effect size interpretations based on van den Berg (2020).

ran a 3*2 RM ANOVA on the proportion of hits and correct rejections with factors probe type (number, upside-down-letter, normal letter) and recall type (cued, free recall).

In general, subjects were more accurate on free recall trials (average ratio of hits and correct rejections per trial was .565, SE_{mean}= .008) than on cued recall trials (.525, SE_{mean}= .005), main effect of recall type ($F(1,35)= 21.758$, $p< .001$, partial $\eta^2= .383$). The effect of probe type was also large and significant ($F(2,34)= 62.381$, $p< .001$, partial $\eta^2= .786$). Subjects decided more accurately on both the number- (.598, SE_{mean}= .008) and upright-letter probes (.589, SE_{mean}= .013; comparison p= .525) than on the upside-down-letter probe (.448, SE_{mean}= .012; both comparisons p< .001). The analysis also indicated a significant interaction effect between probe type and recall type: $F(2,34)= 4.732$, $p= .015$, partial $\eta^2= .218$. The tendency for decisions to be more often correct on free recall trials was present for the three probes, but was stronger for the number probe.

Figure 2

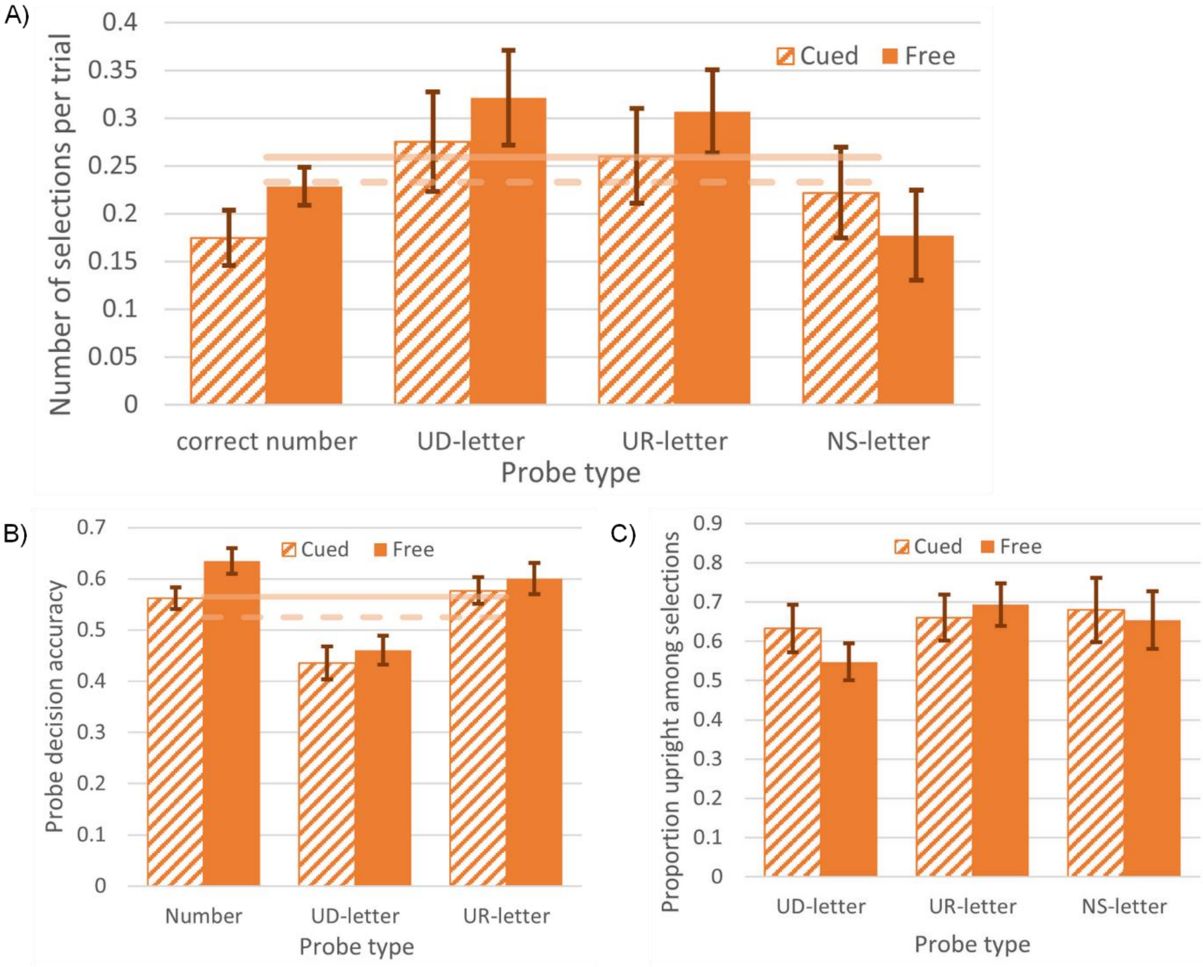
Item Availability in Experiment 1



Note. Figure shows item availability for cued and free recall on blocks 1 and 2. Striped bars and lines indicate cued (partial) recall, full bars and lines indicate free recall. Values reported in de Gardelle et al. (2009) and Sperling (1960) added for reference. Error bars indicate 95% confidence intervals.

Figure 3

Experiment 1 Mean Probe Selections per Trial



Note. Figure compares probe types, separated per trial type (cued, free recall). UD-letter and UR-letter probes refer to probes about the letter shown upside-down and upright in the grid respectively. The NS-letter probe refers to the probe showing a letter that was not shown in the grid. Significant effects not indicated. Error bars indicate 95% confidence intervals. a) Comparison of probe selections per trial. b) Comparison of probes on decision accuracy (hits and correct rejections). c) Comparison of probes on the proportion of selecting its upright alternative.

The previous analyses showed that the special item probes were selected more than the not-shown letter probe, indicating that some information regarding the special items was maintained up to the recognition question. In addition, decisions regarding the upside-down-letter probe were less accurate than those for the number- or normal-letter probes, indicating that the upside-down letter was not remembered as correctly as the upright letter or number. This means that the experimental procedure generated the right conditions for an illusory letter perception to occur. Such illusory perception would be reflected in an increased proportion of upright against to upside-down selections of the upside-down-letter probe, compared to the same proportion for the not-shown-letter probe.

To analyse this, a 3*2 RM ANOVA was performed on the proportion of upright probes among the selected ones, with factors probe type (upside-down letter, normal-letter, and not-shown-letter) and recall type (cued, free). For one person, the proportion could not be calculated because no not-shown-letter probes were selected, so this analysis was based on the 35 remaining subjects.

We observed no effect of recall type ($F(1,34)= 1.158$, p= .289) which means that subjects had similar proportions of selecting the upright alternative of the letter probe types for cued recall (proportion= .658, SE_{mean}= .028) and free recall trials (.632, SE_{mean}= .020). The proportion of upright selected probes did differ strongly between the probe types ($F(2,33)= 6.745$, p= .004, partial η^2= .290). For the upside-down-letter probe (.590, SE_{mean}= .018), the proportion of upright selections was lower than both the upright-letter probe (.677, SE_{mean}= .024; p < .001) and the not-shown-letter probe (.667, SE_{mean}= .033; p= .009). The upright-letter- and not-shown-letter probe types' proportion of upright selections did not differ (p= .685). In other words, subjects generally reported the upright letter and not-shown letter to have been upright more than the upside-down letter.

There was also a large interaction effect between probe type and recall type ($F(2,33)= 4.572$, p= .018, partial η^2= .217, see also Figure 3c). On cued recall trials, the proportion of upright selections did not differ between the probe types ($F(2,33)= 1.484$, p= .241), but they did on free recall trials ($F(2,33)= 8.678$, p< .001, partial η^2= .345). On free recall trials, they reflected the main effect of probe type: the upside-down-letter probe's proportion upright (.548, SE_{mean}= .023) was lower than the upright-letter probe (.694, SE_{mean}= .027; p < .001) and the not-shown-letter probe (.654, SE_{mean}= .036, p = .016), with the proportions being similar for the upright-letter and the not-shown-letter probes (p = .213).

To summarise, on cued recall trials subjects reported the upside down letter and the upright letter to have been upright equally much as they reported a not-shown letter to have been upright. On free recall trials, they reported the upside-down letter to have been upright less strongly than either the upright letter or the not-shown letter. Interestingly, in both conditions all three letter probes were more often selected as upright than upside-down (all ps< .001 except for the free recall upright-letter probe's proportion, which was marginally above .50: p= .080.

All ps two-sided). This confirms that the experimental stimuli and procedure induced a general tendency to report probes as upright for all three letter probes.

Discussion

This experiment's goal was twofold. First, it investigated whether an upside-down letter presented in a Sperling paradigm would be reported as upright due to an illusion in perception or whether it would reflect a response bias. Second, the experiment assessed a possible role of a cue in directing attention away from the iconic memory representation of the upside-down letter.

Before evaluating whether a response bias or an illusion better explains subjects' report of the upside-down letter, several analyses aimed to confirm that information about the upside-down letter was available at the time of the recognition question. The short decay time of iconic memory (e.g. Sperling, 1960; Teeuwen et al., 2021), where the grid's letters are represented after grid offset, could prevent information from the grid from being available up to the recognition question.

The first and main indication of letter representations in IM was the comparison of item availability between cued and free recall trials, which showed no partial superiority effect. Subjects had as many items available for immediate free recall as for partial recall, so the experimental procedure used here could not show that more items were represented in iconic memory than that could be freely recalled. The similarity of partial report performance in this experiment and what was reported in de Gardelle et al. (2009) implies that the 300ms display prevented a partial report superiority to occur.

Perhaps subjects used the 300ms grid display to attentionally select a row or a few items to retain in working memory, effectively self-cuing those items. By maintaining items in working memory, they would be protected from overwriting by the backward mask (e.g. Sligte et al., 2008). Working memory's small capacity would explain the relatively weak partial recall performance found here and that is reported in de Gardelle et al. (2009). The fact that Sperling's original experiments used 50ms display suggests that a shorter display period might prevent subjects from using this strategy.

Even in the absence of a partial report superiority effect in our data, subjects did retain some information regarding the items that had been in the grid. The analysis of probe selection rates showed that upside-down-letter- and upright-letter probes were selected more often than the not-shown-letter probe. This indicates that subjects could differentiate between items that had been part of the preceding grid from those that were not.

The analysis of probe accuracy further indicated that subjects did not accurately remember the grid's special items. Rather, the accuracy of decisions regarding the number probe (56-63%) even suggests that a lot of information had been decayed before the recognition task. The recognition question required subjects to make independent decisions on four different probes would have only contributed to this decay. In spite of this, the performance on the upside-down-letter probe was considerably worse compared to the number- and upright-letter probe types, confirming that the procedure used could show an illusory perception of the upside-down letter.

The proportions of selecting upright over upside-down alternatives did not reveal an illusion effect. For cued recall, the upside-down-letter probe was selected upright as much as the upright-letter probe and the not-shown-letter probe. Subjects judged the upside-down letter's orientation to be similarly upright as the upright letter, both being above the .50 proportion of probed letters being upright. This shows that there is a tendency to report the upside-down and the upright-letter as upright, conceptually replicating de Gardelle et al.'s (2009) finding of illusory perception. However, the inclusion of the not-shown-letter probe showed that the tendency to report upright did not differ from a letter than was not shown in the grid at all. The tendency to report letters as upright is therefore not exclusive to upside-down letters, as supposed by the illusory-perception account, but generalises to letters that were not actually shown in the grid.

On free recall trials, the tendency to select upright probe alternatives was even lower for the upside-down-letter probe than for the not-shown letter probe. This finding suggests that the upside-down letter is represented more veridically than the contextual bias would suggest. Interestingly, the tendency to select upright probe alternatives did not differ between the upright-letter probe and the not-shown letter probe. Possibly this could suggest that subjects did not notice any upside-down letter in the grid: the noticed presence of upside-down letters would likely decrease the prior of a not-remembered letter to have been upright. The high occurrence of upside-down letters during the experiment suggests that this is not likely, but at the same time there was a consistent tendency to report letter probes as upright.

To summarise the results on the first aim of the experiment, it is clear that there was no illusory letter perception. The tendency to select upright probe letters was as great for the upside-down-letter probe as for the not-shown-letter probe. However, it is unclear whether or how the upside-down letter was processed up to the recognition question. There was no partial report superiority effect to indicate the maintenance of items in non-cued rows in iconic memory after grid offset, but subjects did select the upside-down and upright letter probes more than the not-shown letter probe. Some information about the upside-down letter may have been available up to the recognition question. Additionally, it could be possible that subjects focused on item identity while ignoring letter orientation. It would explain the increased selection rates for special item probes while probe selection accuracy was relatively low.

The second aim of this first experiment was to assess the role of the cue in directing attention to IM item representations. Increased selection rates for the special items on free recall trials compared to cued recall trials suggests that the cue does limit processing of special items, which only occurred in non-cued rows. Moreover, the probe selection accuracy was better on free compared to cued recall trials, showing that the increased attention to the special items leads to more accurate decisions on them when they are probed. This improved accuracy was reflected in the proportion of upright probe selections on free recall trials, which was lower for the upside-down-letter probe compared to the not-shown-letter probe baseline.

Attending an upside-down letter during grid presentation therefore seems to make it processed veridically rather than illusorily. A possible explanation for this is the 300ms grid display time. Lamme and Roelfsema (2000) argue that higher visual processing areas can be reached within 100ms after stimulus presentation and that feedback between layers takes approximately 10ms. Within the 300ms that the grid is presented on the screen, the continued sensory input of the upside-down letter might update predictions of that letter being upright before grid offset removes sensory evidence of the letter being upside-down.

A possible illusion effect could therefore still be found if items that are not attended are maintained in iconic memory after a briefer display time. At the same time, an easier probe task than the one used in this experiment should make a possible illusion effect easier to elicit.

Experiment 2

The second experiment attempts to address the issues discussed in the first experiment by shortening the grid presentation time and adapting the recognition question to make it less demanding. A shorter grid presentation time might reintroduce a partial report superiority effect which is indicative of items being retained in iconic memory, prevent subjects from selectively attending specific items during grid presentation, and prevent sustained visual input from updating predictions about the grid's items. The changed recognition question will allow for a more direct comparison between upside-down, upright, and not-shown letters.

Methods

Participants

In total, 49 subjects took part in this experiment. All were part of the university's student pool and were compensated for their participation with credits as part of a course requirement. Two subjects are not included in the analysis because they did not do the task seriously, six subjects were excluded from the initial analysis because they ignored the cue in answering the recall question. The 41 included participants had a mean age of 19.2 years (SD= 1.6) and consisted of 33 women, 4 men, and 4 people who did not identify as either man or woman.

Stimuli and procedure

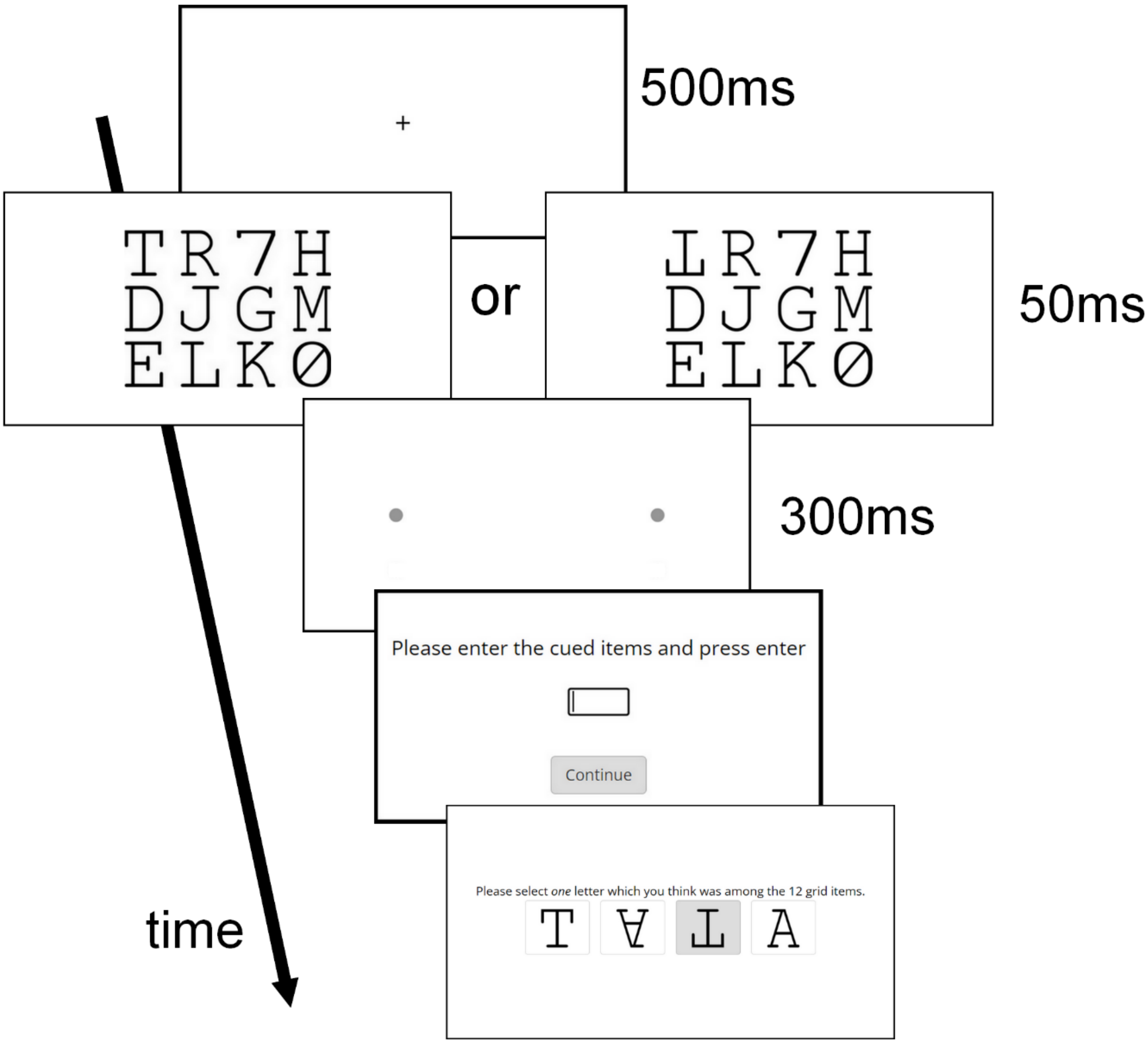
The experiment proceeded similarly to the first, but with a 50ms grid presentation time. Earlier work on the Sperling paradigm suggests that a shorter presentation time improves the partial report superiority effect (Sperling, 1960), possibly by preventing attentional selection of specific items already during grid presentation. Additionally, it should prevent sustained sensory input from updating predictions while the grid is displayed. The grid composition was changed slightly such that it still contained 9 normal letters and three special items, but the three special items were two numbers and the target letter. On every trial, the target letter was either upright or upside-down in the grid, creating two trial types (henceforth UPR and UD, respectively). Because there was not an upside-down letter on every trial, the special items were supplemented with a second number which served to reduce a possible attentional pop-out of the upside-down letter when it did occur. Like in Experiment 1, the three special items never occurred in the row that was cued.

The two trial types (i.e. with an upright or upside-down target letter; UPR and UD) were presented in a randomly interleaved fashion in 4 blocks of 30 trials (15 of each trial type). A general instruction before the blocks started informed subjects of the task, followed by two practice trials to get subjects familiar with the trial presentation. There was no free recall condition in this experiment as the previous experiment showed it only facilitated people to attend the upside-down letter. The similarity of free recall item availability in Experiment 1 was similar to other studies using the Sperling paradigm, suggesting that it is relatively robust to varying task parameters.

We also changed the recognition question to make it easier for subjects to indicate their memory of letter identity and orientation. It now probed the target letter (which had been shown in the grid as upside-down or upright, depending on the trial type) and a letter that was not shown in the preceding grid (foil). Both letters were probed in their upright and upside-down orientations, creating four alternatives: target-upr (upright), target-ud (upside-down), foil-upr, foil-ud (see Figure 4). Subjects selected only one of these, which they thought had been in the preceding grid. This forced every subject to consider both letter identity and orientation in their selection. The probes were presented in a random order.

Figure 4

An example trial in Experiment 2



Note. The target letter in this example trial is the 'T' (top-left in the grid). The upside-down target probe is highlighted in grey in the recognition question. Subject entries on the cued recall and recognition question are self-paced.

Data-analysis

Similar to experiment 1, the letters entered on the recall question were counted correct if they occurred in the cued row, irrespective of their relative placement. Performance on the recall question was calculated by averaging the percentage of cued letters that the subject entered. This percentage was then analysed in a RM ANOVA using trial type (UPR and UD) and block (1 to 4) as factors. A reconstructed availability of more than 4-5 items would be taken to show that more items are represented in IM than could be freely reported.

A potential difference in trial difficulty between the two trial types, for example because of a greater distraction effect from the upside-down target letter in UD trials, might influence probe selections. Such difference in difficulty would be reflected in the accuracy on the recognition question, measured as the percentage of times the correct probe was selected. This accuracy was analysed with condition and block as independent variables to also investigate a possible learning curve over time.

In this experiment, reports of the upside-down letter as upright was again compared against the general tendency to report not-shown letters as upright. To analyse this, the selection rates for the four probe options were analysed in a RM ANOVA with factors trial type (UPR, UD), block (1-4), probe id (target, foil) and probe orientation (upright, upside-down). The crucial comparison was between the selection rate of upright over upside-down target and foil probes in the condition where the target was shown upside-down in the grid (UD). The same comparison for UPR trials served to compare the illusion strength to a letter that was shown upright in the grid and the blocks serve to investigate whether a potential learning curve in accuracy influenced probe selection distributions.

Follow-up analyses were performed if the RM ANOVA on selection rates for the probe options showed a significant main effect. The interaction between probe id and probe orientation was analysed separately for UPR and UD trials independently of a significant three-way interaction between trial type, probe id and probe orientation because of its relevance as a direct test of the illusory perception of upside-down letters. Such illusory effect would show when the ratio of selecting upright over upside-down letters is larger for the target-letter probes than the foil-letter probes on UD trials.

Results

Recall

The reconstructed percentage of available items on the recall question did not differ significantly between UPR and UD trials ($F(1,40)= 0.013$, p= .908) but did clearly improve over blocks ($F(3,38)= 37.485$, p< .001, partial $\eta^2= .747$). Item availability in block 1 (μ= 39.0%, SE_{mean}= 2.08%) was improved in block 2 (μ= 48.6%, SE_{mean}= 2.44%), which in turn was improved in block 3 (μ= 52.6%, SE_{mean}= 2.32%) and superseded by the availability of items in block 4 (μ= 55.8%, SE_{mean}= 2.32%; all ps< .008). The trial type by block interaction was significant ($F(3,38)= 3.067$, p= .039, partial $\eta^2= .195$), but did not show a clearly interpretable pattern (see Figure 5a).

The overall mean availability of 49.0% (SE_{mean}= 2.14%) items (approximately 6 out of 12) was a clear improvement compared to experiment 1, which showed a mean availability of 40.2% of items (SE_{mean}= 2.04%; $t(75)= -2.949$, p= .002). The improvement over blocks could be interpreted as showing a learning curve, which is reflected in the blockwise comparisons with cued recall item availability in Experiment 1. Item availability in block 1 did not differ from that in Experiment 1 ($t(75)= 0.422$, p= .337), but availability in blocks 2 to 4 was better than in Experiment 1 ($t(75)= -2.607$, p= .006; $t(75)= -3.938$, p< .001; $t(75)= -4.976$, p < .001 respectively, all four p-values one-sided). The number of available items in Experiment 2 was still below Sperling's (1960) original findings.

Recognition - Accuracy

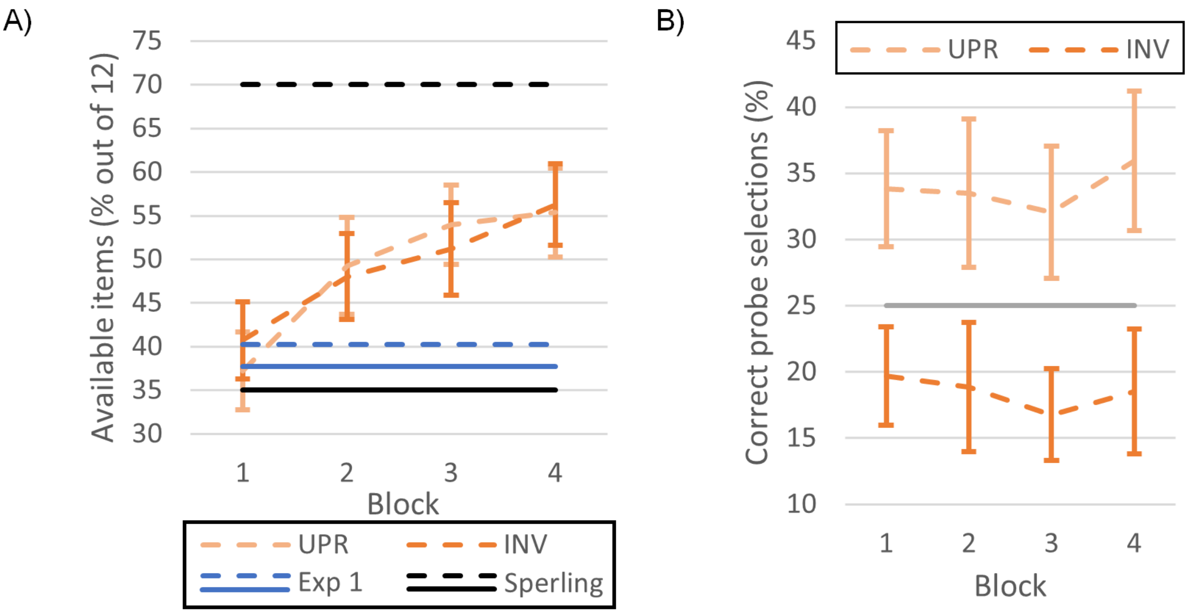
Accuracy on the recognition question strongly differed between UPR and UD trials ($F(1,40)= 24.650$, p< .001, partial $\eta^2= .381$; Figure 5b): on UPR trials, the correct probe was only selected on 33.8% of trials, (SE_{mean}= 1.84%) whereas it was 18.5% (SE_{mean}= 1.52%) on UD trials. The proportion of correct probe selections did not differ between blocks ($F(3,38)= 0.870$, p= .465), nor was there an interaction effect between trial type and blocks ($F(3,38)= 0.282$, p= .838). Probe selection accuracy was above an objective guess rate of 25% for UPR trials on all four blocks (all $t(40)> 2.852$, all ps<.004) and below 25% for UD trials (all $t(40)< -2.537$, all ps< .009).

Recognition – Probe selection distribution

The analysis of probe selection distributions indicated how often target and foil probes were selected as upright. The RM ANOVA on probe selections revealed that all effects including block were not significant: all *F*s< 1.500, all ps> .229, which means that probe selection distributions were similar across blocks. For further analyses in this experiment, probe selection distributions will not be considered separately for blocks.

Figure 5

Performance on Recall (a) and Recognition (b) Questions in Experiment 2



Note. Error bars indicate 95% confidence intervals. Dashed lines indicate cued recall, continued lines indicate free recall. (a) Shows the percentage of available items per block separately for UPR and UD trials. Values for experiment 1 and Sperling are included for reference. (b) Shows the percentage of correct probe selections per block separately for UPR and UD trials. Grey line indicates 25% chance level.

We did find main effects of probe id ($F(1,40)= 6.354$, $p= .016$, partial $\eta^2= .137$) and probe orientation ($F(1,40)= 21.562$, $p< .001$, partial $\eta^2= .350$). Overall, target probes were selected more than foil probes ($\mu= 25.9\%$, $SE_{mean}= 0.35\%$ for target probes, $\mu= 24.1\%$ $SE_{mean}= 0.35\%$ for foil probes) and upright probes were selected more than upside-down probes ($\mu= 30.9\%$ $SE_{mean}= 1.27\%$ and $\mu= 19.1\%$ $SE_{mean}= 1.27\%$, respectively). In other words, there was a tendency to select the target-letter probes over not-shown-letter probes and a tendency to select upright probes over upside-down ones.

A significant interaction between probe id and probe orientation ($F(1,40)= 9.151$, $p= .004$, partial $\eta^2= .186$) revealed that upright- and upside-down probe selections differed between target and foil probes. Upright target probes were selected more often than upright foil probes ($t(40)= 3.613$, $p< .001$; 33.0% $SE_{mean}= 1.51\%$ versus 28.8% $SE_{mean}= 1.27\%$), but selections of upside-down letters did not differ between target and foil probes ($t(40)= -0.685$, $p= .249$; 18.8%

SE_{mean}= 1.41% and 19.4% SE_{mean}= 1.30% respectively). Overall, subjects tended to report target probes as upright more than foil probes.

There was no interaction effect between trial type and probe type ($F(1,40)= 2.758$, $p= .105$), which means that the propensity of subjects to select target probes over foil probes was similar between UPR and UD trials. The tendency to select upright over upside-down probes also did not differ between UPR and UD trials ($F(1,40)= 0.157$, $p= .694$).

Importantly, the three-way interaction between trial type, probe id, and probe orientation was not significant ($F(1,40)= 1.433$, $p= .238$), indicating that the stronger tendency to report the target letter as upright compared to the foil letter did not differ between trials with an upright compared to upside-down target letter. In turn, it suggests that upside-down letters are reported as upright more than not-shown foil letters during the probe question. Given the importance of this result in supporting a possible illusory perception effect, the interaction between probe id and probe orientation was investigated separately for each trial type.

When only probe selections on UPR trials were analysed (the left side of Figure 6), there were strong effects of probe id ($F(1,40)= 7.834$, $p= .008$, partial $\eta^2= .164$) and probe orientation ($F(1,40)= 19.468$, $p< .001$, partial $\eta^2= .327$), as well as a strong interaction effect between probe id and probe orientation ($F(1,40)= 8.132$, $p= .007$, partial $\eta^2= .169$). These replicated the main effects described above: target probes were selected more often than foil probes (26.5% SE_{mean}= 0.52% versus 23.5% SE_{mean}= 0.52%), and upright probes were selected more than upside-down probes (30.8% SE_{mean}= 1.30% versus 19.2% SE_{mean}= 1.30%). Again, upright target probes were selected more than upright foil probes ($t(40)= 3.509$, $p< .001$; 33.8% SE_{mean}= 1.84% for upright target probes against 27.7% SE_{mean}= 1.25% for upright foil probes) whereas upside-down target and foil probes were selected equally often ($t(40)= -0.221$, $p= .413$; 19.1% SE_{mean}= 1.54% for upside-down target probes and 19.4% SE_{mean}= 1.37% for upside-down foil probes). On trials with an upright target letter, subjects tend to identify it correctly among the different probe options.

When only UD trials were analysed (the right side of Figure 6), only a strong general effect of probe orientation was significant ($F(1,40)= 20.453$, $p< .001$, partial $\eta^2= .338$). Upright probes were selected more than upside-down probes (31.0% SE_{mean}= 1.33% per upright probe compared to 19.0% SE_{mean}= 1.33% per upside-down probe). Target probes were not selected more than foil probes ($F(1,40)= 0.370$, $p= .547$), nor did probe id interact with probe orientation

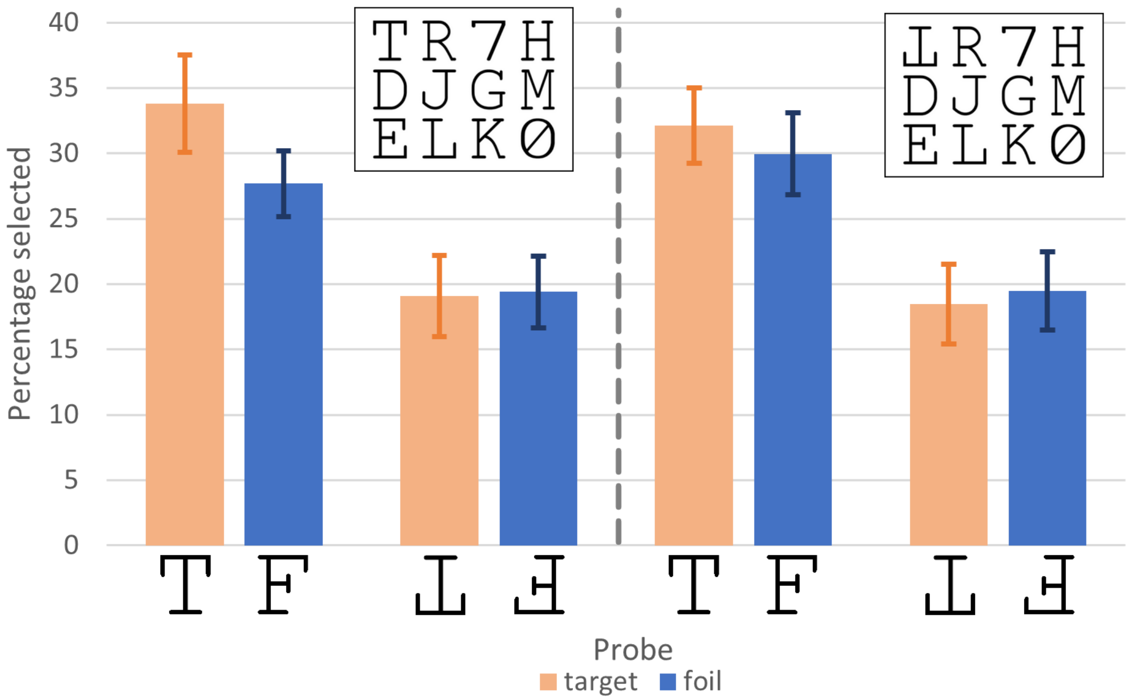
($F(1,40) = 2.759$, $p = .105$). The tendency to select upright probe alternatives did therefore not differ between target and foil probes.

Discussion

The analysis of cued recall answers showed that experiment 2 partially reinstated the cued recall performance expected in a Sperling procedure. A mean performance of 49% correct is clearly above the free recall performance range of around 35% that was replicated in experiment 1, but still below the performance range of 65-75% reported elsewhere (Sperling, 1960). The experiment's online administration may have had a detrimental effect on performance as it puts a stronger demand on instructions, or it could be that the partial report superiority effect is weaker when using visual cues. Black & Barbee (1985) reported that visual cues could also elicit a partial report superiority effect but did note that the effect could be weaker than when using an auditory cue.

Figure 6

Selection Ratios for Recognition Question Alternatives in Experiment 2



Note. This Figure shows how selections were distributed among the probes (individual bars) for UPR (left) and UD trials (right). Error bars indicate 95% CI.

The partial report superiority effect increased throughout the experiment, suggesting that subjects represented more letters in iconic memory as the experiment progressed. It might indicate that partial report is a skill that subjects got better at as they did more trials, although it cannot be excluded that subjects had to get familiar with the task even after the instructions.

The improving availability of items in IM was not paralleled by probe selections distributions, as neither the percentage of correct probe selections nor the distribution of probe selections changed over blocks of trials. The cue is supposed to direct subjects' attention to a specific set of letters that are represented in iconic memory, thereby maintaining their availability. Unattended letters in iconic memory would have decayed independently of the number of letters whose representation in iconic memory was boosted by cue-driven attention, explaining why the probe selection distribution did not change as the experiment proceeded.

The analysis of probe selections showed that subjects were overall able to distinguish target letters from probe letters as they selected target probes more than foil probes. Even though the recognition question concerned non-cued letters, some information about the letters in the grid was therefore still available to subjects when the probes were shown. The main interaction effect between probed letter and probed orientation showed that the general tendency for subjects to report letters as upright was stronger for target probes than foil probes. This tendency to select upright probes could also explain why the probe selection performance was below chance on trials with an upside-down target letter.

There was no three-way interaction effect with trial type, suggesting that this tendency to report target-letter probes as upright more than foil-letter probes applies to upright as well as upside-down target letters, so we repeated the analysis on trials with upright and upside-down targets separately. These showed that the main interaction effect described above was mostly driven by trials with upright target letters, for which the probe selections corresponded to veridical perception. Probe selections on trials with upside-down targets, however, indicated that the proportion of selecting upright over upside-down probes was the same for target and foil probes. The tendency to select upside-down target letters as upright can therefore be explained by a general tendency to report letters as upright.

On trials with an upright target letter, probe selections corresponded to veridical target perception. This showed in the greater number of selected target probes and the stronger tendency to report a letter as upright when it was the target letter compared to when it was the foil letter. On trials with upside-down targets, there was no evidence that subjects remembered the target letter as subjects selected target and foil probes equally often. In addition, the tendency to report upside-down target letters as upright was statistically similar to the tendency to report foil letters as upright, even though upside-down target letters were numerically more often selected as upright than foil letters. The p-value for this effect ($p= .105$) does not allow for a strong rejection of such an illusion effect, however, as an effect could have been obscured by noise. A study with more power should be able to identify such an effect if it exists.

Experiment 3

The second experiment did not provide evidence to suggest that upside-down letters are illusory perceived as upright. The absence of such evidence might be caused by a lack of power, so the current experiment will be a full-power replication of experiment 2.

Methods

Participants

In total, 196 people partook in the experiment. They were recruited using the university's student participant pool and were compensated with credits to partially fulfil a course requirement. Ten participants were rejected because they did not provide serious answers. Fifty-three participants ignored the cue and self-selected a row to enter items from, they were not included in the analysis. The analysis therefore included 134 participants, resulting in >95% power to detect a significant interaction effect between the tendency to report upside-down target letters and foil letters as upright (at $\alpha= .05$ and partial $\eta^2= .065$). Participants had a mean age of 19.8 years ($SD= 2.7$), of which 110 were women, 21 men, and 3 did not identify as either woman or man.

Stimuli and procedure

The experiment was almost an exact replication of Experiment 2. The improvement in recall performance over blocks found in experiment 2 might indicate that subjects were unfamiliar with the procedure despite having two practice trials. The instructions now mentioned the durations of trial sections to prepare subjects for the short grid presentations and subjects could repeat the two practice trials until they felt they were familiar with the trial procedure.

Furthermore, some subjects in Experiment 2 had to be rejected because they ignored the cue when answering the recall question. To avoid subjects from using such strategy, the instructions mentioned that it would be impossible to look at all letters in turn and that prematurely focusing attention on any single row would be counterproductive. During the breaks between blocks, subjects were reminded to rest their gaze on the fixation cross at the start of every trial and to have their attention be directed by the cue. These changes did not have the desired effect since 53 out of 186 participants still used this strategy (28%), whereas only 6 out of 47 (13%) used this strategy in Experiment 2. Potentially, advising against the strategy had an opposite effect on subjects.

Data analysis

Data were analysed identically to Experiment 2.

Results

Recall

The RM ANOVA on available items with factors trial type and block revealed no effect of trial type ($F(1,133)$= 1.914, p= .169). The average availability of items on UPR trials was comparable to that on UD trials (47.0% SE_{mean}= 1.46% and 46.4% SE_{mean}= 1.48%, respectively). There was a strong effect of block ($F(3,131)$= 69.359, p< .001, partial η^2= .614). Each block's item availability was better than the previous one, an effect that was more pronounced for blocks 2 (μ_{block1}= 36.7%, SE_{mean}= 1.35%; μ_{block2}= 46.9%, SE_{mean}= 1.61%) and 3 (μ= 50.9%, SE_{mean}= 1.61%; ps<.001) than block 4 (μ= 52.4%, SE_{mean}= 1.64%; p=.026).

A significant interaction between trial type and block ($F(3,131)$= 3.099, p= .029, partial η^2= .066) shows a larger initial improvement for trials with upright compared to upside-down targets in the first blocks that reversed in later blocks (see Figure 7a). This pattern, suggestive of a plateauing learning curve that is stronger for trials with upright targets, can in hindsight also be seen in experiment 2, albeit not as clearly.

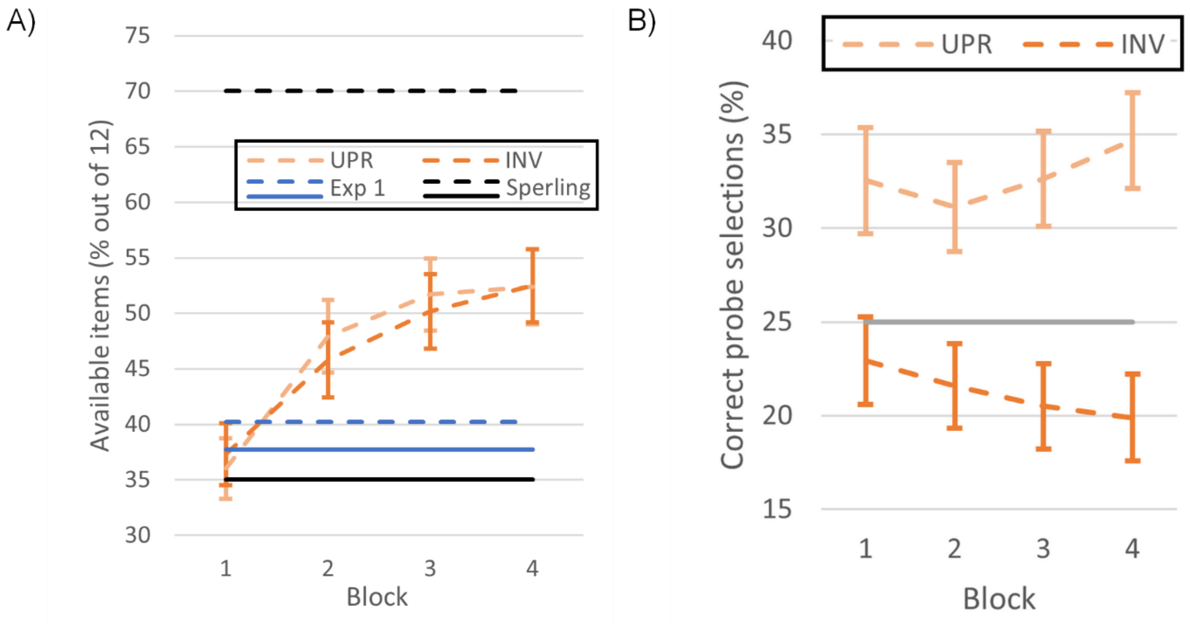
The mean availability of items on the recall question was 46.7% (SE_{mean}= 1.46%), comparable to the availability in Experiment 2 (49.0%; $t(173)$= 0.790, p= .215) and showed a similar development over the four blocks. The mean availability of items did not differ from the mean availability on cued recall trials in Experiment 1 for block 1 ($t(168)$= 1.267, p= .104), but it did for blocks 2, 3, and 4 ($t(168)$= -2.019, p= .023; $t(168)$= -3.265, p< .001; $t(168)$= -3.649, p< .001, all four p-values one-sided).

Recognition - Performance

The RM ANOVA on the percentage of correct probe selections with factors trial type and block showed a clear effect of trial type ($F(1,133)$= 60.580, p< .001, partial η^2= .313), with a better

Figure 7

Experiment 2 Recall and Recognition Performance



Note. Dashed and continuous lines indicate cued and free recall, respectively. Error bars indicate 95% confidence interval. A) Shows the number of items available for immediate recall across blocks. Values for experiment 1 (blue) and in Sperling (1960; black) added for reference. B) Shows the percentage of correct probe selections across blocks. Grey line indicates the 25% objective guess rate.

performance on UPR trials (32.7% correct, SE_{mean}= 0.94%) than UD trials (21.2% correct, SE_{mean}= 0.78%). There was no effect of block ($F(3,131)$= 0.827, p= .481), the interaction between trial type and block was marginally significant ($F(3,131)$= 2.450, p= .066, partial η^2= .053). Although the difference in performance between the trial types numerically increases over blocks, performance levels do not develop for UPR trials ($F(3,131)$= 2.123, p= .100) or INV trials ($F(3,131)$= 1.365, p= .257) separately (see Figure 7b).

Performance on the recognition question was significantly above 25% for UPR trials (p< .001 for all four blocks) and performance on INV trials was significantly below 25% on all four blocks (p_1= .041, p_2= .002, p_3< .001, and p_4< .001; all eight p-values one-sided, see also Figure 7b). In terms of selecting the correct probe, these results match those in Experiment 2.

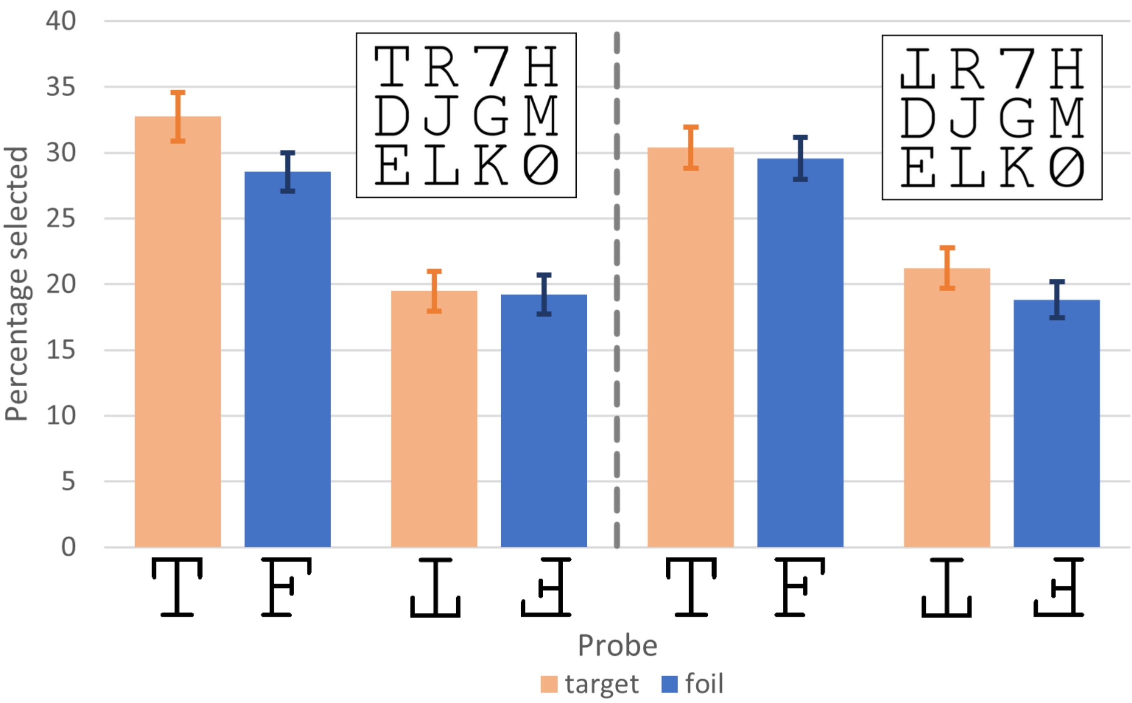
Recognition – probe selection distributions

The RM ANOVA on the distribution of probe selections with factors trial type, block, probe id and probe orientation repeated the tendencies of participants to select target probes over foil probes ($F(1,133)= 18.312$, $p< .001$, partial $\eta^2= .121$; target probes: $\mu= 26.0\%$, $SE_{mean}= 0.23\%$, foil probes: $\mu= 24.0\%$, $SE_{mean}= 0.23\%$) and to select upright probes over upside-down ones ($F(1,133)= 71.014$, $p< .001$, partial $\eta^2= .348$; upright probes: $\mu= 30.3\%$, $SE_{mean}= 0.63\%$, upside-down probes: $\mu= 19.7\%$, $SE_{mean}= 0.63\%$) found in Experiment 2.

Contrary to Experiment 2, the three-way interaction between trial type, probe id and probe orientation was significant ($F(1,133)= 10.972$, $p= .001$, partial $\eta^2= .076$; see Figure 8). It indicates that the target-foil difference in selecting the upright compared to upside down probe differs between UPR and INV trials, as was speculatively suggested in Experiment 2. Interestingly, the general interaction effect between probe id and probe orientation that was significant in Experiment 2 was not significant in this larger dataset: $F(1,133)= 2.008$, $p= .159$.

Figure 8

Selection Ratios on the Recognition Question for UPR and INV Trials in Experiment 3



Note. This Figure shows how selections were distributed among the probes (individual bars) for UPR (left) and UD trials (right). Error bars indicate 95% CI.

The follow-up analysis of probe selection distributions in UPR trials showed an effect of probe id ($F(1,133)$= 12.483, p< .001, partial η^2= .086) and a strong effect of probe orientation ($F(1,133)$= 69.353, p< .001, partial η^2= .343). Target probes were selected slightly more often than foil probes (μ= 26.1%, SE_{mean}= 0.32% and μ= 23.9%, SE_{mean}= 0.32%, respectively) and upright probes were selected much more often than upside-down probes (μ= 30.6%, SE_{mean}= 0.68% and μ= 19.4%, SE_{mean}= 0.68% respectively). The interaction between probe id and probe orientation was also significant ($F(1,133)$= 11.244, p= .001, partial η^2= .078). T-tests revealed that the percentage of upright target probes selections (μ= 32.7%, SE_{mean}= 0.94%) was greater than the percentage of upright foil probe selections (μ= 28.5%, SE_{mean}= 0.74%; $t(133)$= 4.170, p< .001), but that the percentage of upside-down target probe selections (μ= 19.5%, SE_{mean}= 0.76%) was equal to that of upside-down foil probe selections (μ= 19.2%, SE_{mean}= 0.76%). These results are similar to what was found in experiment 2 and indicate that targets (veridically) selected upright more than foils.

The analysis of INV trials showed slightly more selections of target probes (μ= 25.8%, SE_{mean}= 0.30%) than foil probes (μ= 24.2%, SE_{mean}= 0.30%; $F(1,133)$= 7.163, p= .008, partial η^2= .051) and a strong tendency to select upright probes (μ= 30.0%, SE_{mean}= 0.64%) over upside-down probes (μ= 20.0%, SE_{mean}= 0.64%; $F(1,133)$= 59.691, p< .001, partial η^2= .310). The interaction between probe id and probe orientation was not significant ($F(1,133)$= 1.840, p= .177), which means that the tendency to select the upright over upside-down letters was as great for target letters as for foil letters. Numerically, the interaction effect even goes in the direction of veridical perception (i.e. upside-down probe selections are more often target probes than foil probes, when the selected probe is upright it is equally likely to be a target or foil probe).

Discussion

The goal of this experiment was to replicate experiment 2 but with a better power to find a possible effect of upside-down letters being perceived as upright. Such effect would show up as an increased tendency to select upright target probes compared to foil probes on trials where the target letter was shown upside-down in the grid. Like in experiment 2, no such effect was found even though with the sample size there was a 95% power for it to be detected.

The three-way interaction between trial type, probe id and probe orientation (p= .001) showed that probe selection distributions differed between trials with an upright and upside-down target letter. When the target letter was upright, subjects selected target-letter probes more than foil-letter probes, showing they retained some information about which letters were shown in the grid. They also demonstrated to have retained some information about the target letter's orientation, as indicated by a greater tendency to select upright probes when they showed the target letter than when they showed a foil letter. This shows that it was possible for subjects to retain identity and orientation information regarding a non-cued letter when it was normally oriented.

When the target letter was upside-down, subjects demonstrated to be able to distinguish the shown probed letter from the not-shown one. At the same time, there was no evidence that they retained information about its orientation as the upright/upside-down selection rates were similar for the target- and foil-letter probes. As the most direct test of illusory perception, the similarity in selecting upright and upside-down probe alternatives between target and foil probes does not show that the upside-down letter is perceived as upright.

Statistically, then, there is a clear dissociation between the information retained about upright target letters and upside-down target letters. For upright letters, subjects demonstrate to have retained information about both the target's identity and orientation, but for upside-down letters subjects only retain information about the letter's identity. Like in Experiments 1 and 2, it replicates the tendency to report upside-down letters as upright, but again this tendency is equally strong for letters that had not been shown in the grid in the first place. The results shown here therefore reject the idea that upside-down letters are illusorily (and hence selectively) reported as upright.

Apart from the distribution of probe selections, a secondary question that this experiment aimed to investigate was the improvement over blocks. A possible explanation for it was that subjects might not have been familiar with the trial presentation or task demands, which was addressed by slightly altering the instructions. Subjects were informed more explicitly about the trial timings and they were able to repeat practice trials at will.

Despite these changes, there was still a clear pattern of improvement in recall performance as the experiment progressed. The numerically decreasing improvement over blocks suggested that performance might have approached a ceiling level, although the present study could not formally test this. Interestingly, such a ceiling would still be below Sperling's (1960) partial report performance. The development of recall performance over blocks also differed between trials with upright and upside-down target letters. Visually, the (learning) curve for trials with upright targets approached its maximum level faster than the curve for trials with upside-down targets, possibly reflecting different learning rates between the two trial types.

Like in Experiment 2, the improvement in cued recall performance was not reflected in performance on the recognition question. The correct probe was selected more than chance would predict on trials with upright targets but less than chance when the target was upside-down. The performance did not change over time, again indicating that the decay of unattended items in iconic memory is independent of the number of items that is represented. The significantly below-chance performance when the target is upside down is again likely the result of a general tendency to select upright probes. The above-chance performance on trials with upright targets could be attributed to the general tendency to select upright probes, combined with a general recognition of the target letter as revealed by the larger selection rate for target probes compared to foil probes.

General discussion

The experiments presented here investigate the reported phenomenon where letters presented upside-down can be illusorily perceived as upright. The results in Experiment 1 replicated the tendency to report letters as upright, but found that this tendency is not exclusive to letters that were actually displayed upside-down. The 'illusory perception' of letters as being upright was found to extend to letters that were not shown at all. Additionally, we found that the procedure used in de Gardelle et al. (2009) could not provide evidence that letters were represented in iconic memory, where the influence of predictions on perceptual processing was hypothesized to occur, in the first place. Experiments 2 and 3 confirmed that even when items were maintained in iconic memory, the tendency to report letters as upright applies to letters that were not shown as well as shown upside-down letters.

Iconic memory

Evidence for a large-capacity iconic memory, as indicated by a greater availability of items in partial report compared to free report, is important in defending a PP account of illusory letter perception. When letters are represented in iconic memory without continued bottom-up sensory input, the precision of the bottom-up signal will weaken. Top-down predictions will likely have an increasingly stronger influence on the representation or processing of the visual input (de Lange et al., 2018).

The low partial report performance found in Experiment 1 did not demonstrate that subjects retained more items in iconic memory than they could report. Remembered letters would likely be the ones that the subject happened to attend to, but this attention only made the representation of such letter more accurate. As discussed in Experiment 1, subjects may not have encoded unattended items in iconic memory beyond their presentation on screen at all. It is unlikely that top-down influences would be able to modulate the relatively strong bottom-up signal to generate an illusory upright perception of an upside-down letter. Indeed, the tendency to report the upside-down letter as upright did not exceed the same tendency for a letter that was not shown. Importantly, the partial report performance in Experiment 1 was similar to the performance reported in de Gardelle et al. (2009). As such, it seems likely that their finding that upside-down letters are more often reported as upright is based on a tendency to generally report letters as upright, which would undermine the account of illusory perception.

In Experiments 2 and 3, performance on partial report indicated that the number of letters represented in iconic memory superseded the number that could be reported. Although the number of items that are available immediately after grid display (5-6 out of 12) was not as great as reported elsewhere (8-9 out of 12; Sperling, 1960), it was still significantly above the items available for free report. The probable cause for this change is the shortened grid presentation time from 300 to 50ms, which is consistent with the idea that a different attentional strategy removes the partial report superiority effect when the grid is shown for longer. Additionally, the improvement in partial report performance across blocks allows for speculation that the representation of information in iconic memory and/or retrieval of information from it can be learned. Future studies should address this possibility more directly, however.

Uncertainty rather than perception drives tendency to report upright

The partial report superiority found in Experiments 2 and 3 should have made the illusory upright perception of upside-down letters more likely to occur. Still, the results showed no convincing evidence for such illusory perception. There was a tendency to report the upside-down letter as upright, but like in Experiment 1 this tendency reflected a general bias towards reporting letter stimuli as upright rather than as upside-down. Importantly, the fact that an upright letter was identified more and selected as upright more than a not-shown letter shows that the paradigm would have been able to find a similar effect for upside-down letters if it did exist. Rather than an illusory perception of the upside-down letter, the tendency to report letters as upright seems to originate in post-perceptual processing.

Indeed, there is evidence that letter or textual recognition can occur at an abstract level (i.e. independent of a visual representation), avoiding a necessary mental rotation of letters to an upright orientation before they can be recognised as the letters they are (Perea et al., 2018;

Witzel et al., 2011; Yang et al., 2019). The illusory letter perception account implicitly supports such mental rotation as it supposes that upside-down letters are rotated to an upright orientation already during perceptual processes, which makes them recognisable as a specific letter. At the same time, perceptual learning accounts of letter recognition propose that upright letters are more easily recognised than unusually oriented letters (Dehaene et al., 2005). In PP terms, letter recognition processes 'predict' upright letters but will also recognise the same letters in other orientations.

Such an account of letter recognition explains the differential results for upright and upside-down letters. When the processing of upright- and upside-down letters' visual input reaches the letter recognition stage, the upright letter is recognised comparatively easily as it aligns with the prediction of letters being upright. There is even evidence that upright letters are represented more strongly in early visual processing when linguistic stimuli are expected (Chang et al., 2015; Heilbron et al., 2020), which would only further facilitate the recognition of upright letters over upside-down ones.

The integration of top-down predictions of upright letters with relatively strong bottom-up signals corresponding to upright letters results in a more precise representation of the letter as upright compared to not-shown letters, for which there is no bottom-up signal to improve the specificity of predicted letter orientations. As input to decision-level processes regarding the probed letters, there would have been relatively strong bottom-up evidence for the target letter as upright. Hence, upright target letters are chosen more than foil letters and they are selected as upright more than the foil letter.

For upside-down letters, the unpredicted and perhaps weaker upside-down letter activity (compared to upright letters) would lead to less certainty in the identification of the letter. As input to decision-level processes, the weak information regarding the target letter combined with the qualitative difference between target and distractor letters would have allowed subjects to select the target letter. The uncertain letter identification would not provide evidence for the letter being in a specific orientation, increasing the influence of top-down predictions of letters generally being upright. Therefore, the orientation of selected target probes would be similar to that of distractor-letter probes, corresponding to the results found in the experiments described here.

Further research could elaborate on the suggested uncertainty regarding the upside-down letter's identification, although important obstacles will have be solved first. A perceptual account of reporting upside-down letters as upright could in theory be tested by directly asking subjects if the display had contained any upside-down letters. However, subjects might start to attend to any irregularly oriented letters, reducing the tendency to report such letters as upright.

Additionally, one could in principle investigate if the uncertainty of identifying upside-down letters is greater than that for identifying upright letters. Such investigations would require such a metacognitive property to be globally available (i.e. available for the subject to report), while it might be a local property of bottom-up and top-down information integration that is not necessarily available for report.

As an investigation of an effect with a supposed origin in iconic memory, a precondition for upside-down letters to be reported as upright was that items were represented in iconic memory in the first place. As a consequence, the target-letter recognition question was asked after ascertaining that the partial report was superior to free report. Any informational processing regarding the (unattended) target letter would have decayed significantly at the time the letter probes were shown. Given that the representation of items in iconic memory in this paradigm is now supported, future research could consider reversing the question order to obtain stronger effects. However, subjects would have to be prevented from attending to or looking for upside-down letters.

Implications for the partial awareness hypothesis

The current results do not challenge the partial awareness hypothesis (Kouider et al., 2010), but do have implications for the debate on the richness of visual experience. Proponents of rich visual experience claim that conscious visual experience overflows what can be reported about that experience (Block, 2011), whereas opponents suggest that the experience of rich vision is driven by relatively poor visual information being enhanced by top-down predictions (Kouider et al., 2010).

Proponents and opponents of the partial awareness hypothesis have opposing predictions with respect to the experience of upside-down letters. Block would argue that they are (veridically) experienced as upside-down, whereas Kouider et al. (2010) claimed that it would be consciously experienced as upright after top-down predictions influenced representations of the upside-down letter to being upright. The results shown here reject the idea that predictions have such influences. Consequently, there is no evidence for an illusion to separate predictions of Block's idea of perceptual overflow from those of Kouider et al.'s (2010) partial awareness hypothesis in this paradigm: bottom-up processing as well as top-down-influenced processing of letters would both result in the same letter representations.

Another way to find evidence for a top-down-driven change in perceptual processing is to analyse the neural representation of ostensibly illusorily perceived letters. A multivariate pattern analysis could identify typical neural responses to upright and upside-down letters, which can be compared against the neural response to an upside-down letter that is reported as upright. An analysis of the temporal and/or spatial correspondence of the 'illusorily' perceived letter will reveal whether the neural representation will correspond to an upright letter and where in the visual processing hierarchy such representational rotation occurs if it does. Current methods to elicit upright reports of upside-down letters require the simultaneous presentation of other letters, increasing the visual complexity and reducing the signal-to-noise ratio for investigating the brain's response to the letter of interest. Further developments in neuroimaging methods might make such analyses more feasible in the future.

Summary and conclusion

The experiments presented here provide strong evidence against the idea that upside-down letters are illusorily represented as upright in iconic memory. In the first experiment, it was shown that a relatively long display of letters did not result in recall performances to indicate that letters were represented in iconic memory in the first place. Regardless, the upside-down letter was judged equally as upright as a not-shown control letter.

The second and third experiment showed that even when recall performance did indicate sustained letter representations in iconic memory, there was no illusory perception. Rather, the upside-down letter was again judged as upright as a not-shown control letter. In contrast, an upright letter was reported significantly more often as upright than a not-shown control letter was. These results indicate that the upside-down letter is not perceived as upright, but that it results in a more uncertain judgment during the letter recognition task. Presumably, this is due to the unexpected orientation that makes the letter harder to recognise. Future research can proceed on these findings in two directions, either by proceeding on the uncertainty account explained here or by investigating the illusory perception account more directly.

The interpretation of letter reports as showing illusory letter perception has been used to argue against the richness of visual experience. The current work reinterprets such findings in terms of a post-perceptual effect, reinvigorating the debate of whether visual experience really is richer than what we can report or whether the richness is an inference based on the influence of predictions on visual processing.

References

Anton-Erxleben, K., & Carrasco, M. (2013). Attentional enhancement of spatial resolution: Linking behavioural and neurophysiological evidence. *Nature Reviews Neuroscience*, *14*(3), 188-200. https://doi.org/10.1038/nrn3443

Becker, M. W., Miller, J. R., & Liu, T. (2013). A severe capacity limit in the consolidation of orientation information into visual short-term memory. *Attention, Perception, & Psychophysics*, *75*(3), 415-425. https://doi.org/10.3758/s13414-012-0410-0

Benvenuti, G., Chen, Y., Ramakrishnan, C., Deisseroth, K., Geisler, W. S., & Seidemann, E. (2018). Scale-invariant visual capabilities explained by topographic representations of luminance and texture in primate V1. *Neuron*, *100*(6), 1504-1512. https://doi.org/10.1016/j.neuron.2018.10.020

Black, I. L., & Barbee, J. G. (1985). Effect of a partial report visual cue on information available in tachistoscopic presentations of alphanumeric characters varying in number and type. *Perceptual and Motor Skills*, *61*(3), 815-820. https://doi.org/10.2466/pms.1985.61.3.815

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences*, *15*(12), 567-575. https://doi.org/10.1016/j.tics.2011.11.001

Brandman, T., & Peelen, M. V. (2017). Interaction between scene and object processing revealed by human fMRI and MEG decoding. *Journal of Neuroscience*, *37*(32), 7700-7710. https://doi.org/10.1523/JNEUROSCI.0582-17.2017

Burns, E. J., Arnold, T., & Bukach, C. M. (2019). P-curving the fusiform face area: Meta analyses support the expertise hypothesis. *Neuroscience & Biobehavioral Reviews*, *104*, 209-221. https://doi.org/10.1016/j.neubiorev.2019.07.003

Chang, C. H., Pallier, C., Wu, D. H., Nakamura, K., Jobert, A., Kuo, W. J., & Dehaene, S. (2015). Adaptation of the human visual system to the statistics of letters and line configuration. *NeuroImage*, *120*, 428-440. https://doi.org/10.1016/j.neuroimage.2015.07.028

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181-204. https://doi.org/10.1017/S0140525X12000477 de Gardelle, V., Sackur, J., & Kouider, S. (2009). Perceptual illusions in brief visual presentations. *Consciousness and Cognition*, *18*(3), 569-577. https://doi.org/10.1016/j.tics.2010.04.006

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How do expectations shape perception? *Trends in Cognitive Sciences*, *22*(9), 764-779. https://doi.org/10.1016/j.tics.2018.06.002

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences*, *9*(7), 335-341. https://doi.org/10.1016/j.tics.2005.05.004

Dick, A. O. (1969). Relations between the sensory register and short-term storage in tachistoscopic recognition. *Journal of Experimental Psychology*, *82*(2), 279-284. https://doi.org/10.1037/h0028147

El-Shamayleh, Y., Kumbhani, R. D., Dhruv, N. T., & Movshon, J. A. (2013). Visual response properties of V1 neurons projecting to V2 in macaque. *Journal of Neuroscience*, *33*(42), 16594-16605. https://doi.org/10.1523/JNEUROSCI.2753-13.2013

Friston, K. (2018). Does predictive coding have a future? *Nature Neuroscience*, *21*(8), 1019-1021. https://doi.org/10.1038/s41593-018-0200-7

Hegdé, J., & Felleman, D. J. (2007). Reappraising the functional implications of the primate visual anatomical hierarchy. *The Neuroscientist*, *13*(5), 416-421. https://doi.org/10.1177/1073858407305201

Heilbron, M., Richter, D., Ekman, M., Hagoort, P., & de Lange, F. P. (2020). Word contexts enhance the neural representation of individual letters in early visual cortex. *Nature Communications*, *11*(1), 1-11. https://doi.org/10.1038/s41467-019-13996-4

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476), 2109-2128. https://doi.org/10.1098/rstb.2006.1934

Kouider, S., & Dupoux, E. (2004). Partial awareness creates the “illusion” of subliminal semantic priming. *Psychological Science*, *15*(2), 75-81. https://doi.org/10.1111/j.0963-7214.2004.01502001.x

Kouider, S., de Gardelle, V., Sackur, J., & Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences*, *14*(7), 301-307. https://doi.org/10.1016/j.tics.2010.04.006

Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*(11), 571-579. https://doi.org/10.1016/S0166-2236(00)01657-X

Öğmen, H., & Herzog, M. H. (2016). A new conceptualization of human visual sensory memory. *Frontiers in Psychology*, *7*, Article 830. https://doi.org/10.3389/fpsyg.2016.00830

Perea, M., Marcet, A., & Fernández-López, M. (2018). Does letter rotation slow down orthographic processing in word recognition? *Psychonomic Bulletin & Review*, *25*(6), 2295-2300. https://doi.org/10.3758/s13423-017-1428-z

Pinto, Y., Sligte, I. G., Shapiro, K. L., & Lamme, V. A. (2013). Fragile visual short-term memory is an object-based and location-specific store. *Psychonomic Bulletin & Review*, *20*(4), 732-739. https://doi.org/10.3758/s13423-013-0393-4

Quilty-Dunn, J. (2020). Is iconic memory iconic? *Philosophy and Phenomenological Research*, *101*(3), 660-682. https://doi.org/10.1111/phpr.12625

Rensink, R. A. (2014). Limits to the usability of iconic memory. *Frontiers in Psychology*, *5*, Article 971. https://doi.org/10.3389/fpsyg.2014.00971

Sligte, I. G., Scholte, H. S., & Lamme, V. A. (2008). Are there multiple visual short-term memory stores? *PLOS One*, *3*(2), Article e1699. https://doi.org/10.1371/journal.pone.0001699

Sligte, I. G., Vandenbroucke, A. R., Scholte, H. S., & Lamme, V. (2010). Detailed sensory memory, sloppy working memory. *Frontiers in Psychology*, *1*, Article 175. https://doi.org/10.3389/fpsyg.2010.00175

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1-29. https://doi.org/10.1037/h0093759

Spratling, M. W. (2010). Predictive coding as a model of response properties in cortical area V1. *Journal of Neuroscience*, *30*(9), 3531-3543. https://doi.org/10.1523/JNEUROSCI.4911-09.2010

Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews Neuroscience*, *15*(11), 745-756. https://doi.org/10.1038/nrn3838

Teeuwen, R. R., Wacongne, C., Schnabel, U. H., Self, M. W., & Roelfsema, P. R. (2021). A neuronal basis of iconic memory in macaque primary visual cortex. *Current Biology*, *31*(24), 5401-5414. https://doi.org/10.1016/j.cub.2021.09.052

van den Berg, R. G. (2020). *Effect size – A quick guide*. SPSS tutorials. https://www.spss-tutorials.com/effect-size/

Walsh, K. S., McGovern, D. P., Clark, A., & O'Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Sciences*, *1464*(1), 242-268. https://doi.org/10.1111/nyas.14321

Witzel, N., Qiao, X., & Forster, K. (2011). Transposed letter priming with horizontal and vertical text in Japanese and English readers. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(3), 914-920. https://doi.org/10.1037/a0022194

Yang, H., Chen, J., Spinelli, G., & Lupker, S. J. (2019). The impact of text orientation on form priming effects in four-character Chinese words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(8), 1511. https://doi.org/10.1037/xlm0000655