Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale
2015, Vol. 69, No. 4, 283–296

1196-1961/15/$12.00 http://dx.doi.org/10.1037/cep0000058

Incidental Learning of Probability Information Is Differentially Affected by the Type of Visual Working Memory Representation

Amanda E. van Lamsweerde and Melissa R. Beck
Louisiana State University

In this study, we investigated whether the ability to learn probability information is affected by the type of representation held in visual working memory. Across 4 experiments, participants detected changes to displays of coloured shapes. While participants detected changes in 1 dimension (e.g., colour), a feature from a second, nonchanging dimension (e.g., shape) predicted which object was most likely to change. In Experiments 1 and 3, items could be grouped by similarity in the changing dimension across items (e.g., colours and shapes were repeated in the display), while in Experiments 2 and 4 items could not be grouped by similarity (all features were unique). Probability information from the predictive dimension was learned and used to increase performance, but only when all of the features within a display were unique (Experiments 2 and 4). When it was possible to group by feature similarity in the changing dimension (e.g., 2 blue objects appeared within an array), participants were unable to learn probability information and use it to improve performance (Experiments 1 and 3). The results suggest that probability information can be learned in a dimension that is not explicitly task-relevant, but only when the probability information is represented with the changing dimension in visual working memory.

Keywords: incidental learning, visual working memory, features, objects, attention

The amount of dynamic information in the world greatly exceeds the capacity to attend to and represent this information. However, due to efficient voluntary and/or involuntary allocation of resources, many tasks can be performed despite these cognitive limits. For example, when driving a car, the position of many other cars around the driver, the colour of the traffic lights, and the possible presence of pedestrians on the side of the road must be monitored, which requires the recruitment of visual working memory (VWM), a short-term store of visual information limited in capacity to about three to four units of information (Alvarez & Cavanagh, 2004; Luck & Vogel, 1997). Although there is far more visual information available than can be represented in VWM at any given time, some parts of the visual world are very likely to change (e.g., a traffic light is likely to change from red to green) whereas other parts are less likely to change (e.g., a stop sign is unlikely to change colour). Therefore, selectively attending to and storing objects or features that are likely to change is an efficient strategy to improve the ability to detect changes in the visual environment. This selection can be guided by explicit task instructions, or by probability information that has been learned incidentally (Beck, Angelone, & Levin, 2004; Beck, Angelone, Levin, Peterson, & Varakin, 2008; Droll, Gigone, & Hayhoe, 2007; van Lamsweerde & Beck, 2011). We examined how the ability to incidentally learn and use probability information to efficiently allocate VWM resources is affected by whether this information is likely to be initially maintained in VWM.

This article was published Online First May 25, 2015.
Amanda E. van Lamsweerde and Melissa R. Beck, Department of Psychology, Louisiana State University.
Amanda E. van Lamsweerde is now at the Department of Psychology, North Dakota State University.
Correspondence concerning this article should be addressed to Amanda E. van Lamsweerde, North Dakota State University, Minard 232, 1210 Albercht Boulevard, Fargo, ND 58102. E-mail: amanda.vanlamsweerde@ndsu.edu

Grouping in Visual Working Memory

Although capacity in VWM is very limited, capacity can be maximized by combining multiple pieces of information into a single unit. For example, a red ball is composed of several spatially connected features from different dimensions: colour, shape, size, texture, and so forth. One way of maximizing VWM capacity is to represent all of these connected features together in a single representation (e.g., a "red ball": Alvarez & Cavanagh, 2004; Luck & Vogel, 1997; Luria & Vogel, 2011; Vogel, Woodman, & Luck, 2001).

However, while the spatial connection of features is a useful perceptual cue for grouping together two features (Xu, 2006), gestalt grouping principles such as proximity (Woodman, Vecera, & Luck, 2003), similarity (Peterson & Berryhill, 2013), or closure (Anderson, Vogel, & Awh, 2013) may also be used to group together multiple features in VWM. For example, in a change detection task—in which participants determine whether a change occurs between a memory display of items followed by a test display—performance is higher when the memory display contains multiple squares of the same colour (similarity-grouping cue) than when all of the colours are unique. This suggests that identical colours are grouped together into a single unit in VWM (Peterson & Berryhill, 2013). Furthermore, Anderson et al. (2013) found that contralateral delay activity (an ERP marker of the number of items in VWM: Vogel & Machizawa, 2004) decreased when closure cues were present. These data suggest that features can be grouped together in VWM via gestalt grouping cues.

What about when two conflicting grouping cues are available? For example, a display with a red oval and a red triangle could potentially be maintained in two ways: as a "red oval" and a "red triangle" (using the spatial connection cue) or a grouping of "red" information (using the feature similarity cue), resulting in separate or no representation for the shape information. If both the colour and shape are task-relevant, then the spatial connection cue would be the most efficient use of VWM capacity, because all of the features can be remembered in two "coloured shape" VWM representations (assuming a capacity of at least two units; Luck & Vogel, 1997). However, if only one dimension is needed to complete a task (e.g., only colour is task-relevant), both "red"s can be remembered together in one feature-similarity representation, rather than two separate connection-based representations. In this case, feature similarity grouping is a better use of VWM capacity.

However, grouping based on feature similarity (e.g., group red items) in the task-relevant dimension (colour) would only occur for these multidimensional objects if participants are able to ignore the task-irrelevant dimension (e.g., shape), which research suggests is possible (Kondo & Saiki, 2012; Wegener, Ehn, Aurich, Galashan, & Kreiter, 2008; Woodman & Vogel, 2008). For example, memory encoding is faster and contralateral delay activity amplitude is reduced when remembering only the colour of colour-orientation or colour-shape items (Woodman & Vogel, 2008). Furthermore, feature changes are more likely to be detected to dimensions that are needed to complete a primary attention task (Droll, Hayhoe, Triesch, & Sullivan, 2005) or to dimensions that are more likely to change in a change detection task (van Lamsweerde & Beck, 2011), even when participants are not instructed about these probabilities. This suggests that the preferential encoding of task-relevant dimensions is possible and can be learned incidentally. This would suggest that when detecting only colour changes, for example, it is possible to selectively encode and maintain only the colours of stimuli. Therefore, if only colour is task-relevant, participants should maximize VWM capacity by remembering the items as colour groupings if there are multiple items of the same colour (Peterson & Berryhill, 2013). While this would effectively increase colour change detection performance, this strategy would preclude storage of dimensions that are not task-relevant.

However, if all of the task-relevant features in a display are unique, then similarity-groupings are not possible. Without similarity-grouping cues, spatial connection would likely to be used to store all connected features together (e.g., red triangle, Luck & Vogel, 1997). While performance in this case would be lower overall (because fewer task-relevant feature would be stored), features in the task-irrelevant dimension should be stored along with the features in the task-relevant dimension. In this study, we were interested in examining how these structural aspects of the unit of representation in VWM (e.g., similarity groupings or connection groupings) may influence how VWM is functionally used in the incidental learning of probability information.

Incidental Learning

Probability information can be learned incidentally (learned without the intention to do so) and used to direct attention and working memory resources. For example, participants can incidentally learn which items tend to co-occur (Fiser & Aslin, 2002), or the probable location of a target in the context of visual search (*contextual cueing;* Chun & Jiang, 1998). Furthermore, probability information can be incidentally learned during the course of a change detection task, and then used to improve performance (Beck et al., 2008; Brady, Konkle, & Alvarez, 2009; Logie, Brockmole, & Vandenbroucke, 2009; van Lamsweerde & Beck, 2011). Frequently co-occurring features are remembered better in a change detection task than randomly assigned feature combinations (Brady, Konkle, & Alvarez, 2009; Logie et al., 2009), which suggests that co-occurring items can be stored together in a single VWM representation (Brady et al., 2009). In addition, it is possible to incidentally learn feature dimensions are most likely to change, and bias VWM toward that dimension. In van Lamsweerde and Beck (2011), objects would change location, colour, or shape, while one change type (e.g., shape changes) occurred more often than the others (75% of the trials). Performance was higher for the dimension that changed most often, indicating that participants had learned the regularities and used them to bias VWM toward that dimension. This suggests that it is possible to learn regularities about what is likely to change and use this information to maintain only the most informative parts of the items.

However, Beck et al. (2008) demonstrated that participants were not able to incidentally learn probability information in a change detection task when the regularities were from a nonchanging dimension (i.e., shape predicted which item would change colour). After a learning phase in which the object that changed colour was always the same shape (100% of the time), participants were no better at detecting colour changes to the predictive shape than any other shape. If participants had learned that one shape was predictive, they could have improved change detection performance by focusing only on the objects that contained the predictive shape. However, in an additional experiment on intentional learning, when participants were instructed that one of the shapes would change colour most often (but were not told which shape), performance was improved for the predictive shape. Therefore, the deficit was in the ability to incidentally learn the regularity, not in the ability to use shape regularities to improve colour change detection.

The findings from Beck et al. (2008) suggest that incidental learning does not occur when probability information is carried in a nonchanging dimension. However, this may have been the case because the predictive dimension was not stored in VWM, which precluded the probability information from being learned. Specifically, in Beck et al. (2008), only a single dimension was explicitly task-relevant (e.g., there were only colour changes), and features within the memory display were repeated (e.g., two blue items could appear). Therefore, the most efficient use of VWM capacity would be to group together items that had identical features (similarity groupings) at the expense of the predictive dimension (shape; Huang & Pashler, 2007). That is, similarity grouping based on the changing dimension (colour) would prevent probability information from the predictive dimension (shape) from being stored in VWM. However, in order for incidental learning to occur, the items or features that carry probability information must be attended (Baker, Olson, & Behrmann, 2004; Jiang & Chun, 2001; Jiménez & Méndez, 1999; Turke-Browne, Jungé, & Scholl, 2005) and the expression of learning depends on the availability of VWM resources (Manginelli, Geringswald, & Pollmann, 2012; Manginelli, Langer, Klose, & Pollmann, 2013). Therefore, we propose that incidental learning likely does not occur if the predictive dimension is not encoded into the VWM representation.

The Current Study

In the current study, we investigated whether the grouping cues used by VWM influence whether probability information from the predictive (nonchanging) dimension is learned and used. We encouraged similarity-grouping VWM representations in Experiments 1 and 3 by using displays that contained repetitions of feature values in the changing dimension in the memory display. In Experiments 2 and 4, connection-grouping representations were encouraged by using displays that contained all unique features. While the similarity-grouping representations should contain the changing dimension only, the connection-grouping representations should contain both the changing and predictive dimensions. Therefore, learning and use of the statistical regularities should only occur when connection-groupings are encouraged (Experiments 2 and 4), and not when similarity groupings are encouraged (Experiments 1 and 3).

In all four experiments, participants detected changes in a single dimension (colour or shape, counterbalanced across participants). The probability information was carried in the nonchanging dimension (i.e., the *predictive dimension*). That is, if participants detected colour changes, the shape dimension was predictive, and vice versa. For each predictive dimension, two features were selected to be the *predictive features*, counterbalanced across participants. For example, when colour was the predictive dimension, for some (randomly assigned) participants, red objects changed shape most often, and for other participants, blue objects changed shape most often. Similarly, when shape was the predictive dimension, one shape (Shape 1) changed colour most often for some participants a different shape (Shape 2) changed colour most often. In the first three blocks of high-weighted trials, 75% of the trials were *predictable trials*, in which the change always occurred to a *predictive object* (any object that contained the predictive feature). For the remaining 25% of the trials (*unpredictable trials),* any randomly selected object that did not contain the predictive feature would change. These blocks were followed by a block of low-weighted trials, in which changes occurred to a predictive object on only 50% of all trials. Changes in performance over the three high-weighted blocks and any change in performance from the high-weighted blocks to the low-weighted block were used to examine how rapidly participants were able to learn and adjust to changes in the probability information. For the high-weighted blocks, we predicted that there would be a probability effect (performance would be higher on the predictable trials that the unpredictable trials) only when connection-groupings were encouraged (that is, when all feature values of the changing dimension within a memory display were unique). For the low-weighted block, the probability effect should persist, demonstrating that the effect is resistant to immediate changes in the probability information.

Experiment 1

Experiment 1 was a near replication of Beck et al. (2008), and was used to test whether probability information is learned when similarity groupings are encouraged. Therefore, in Experiment 1, at least two objects within the memory display shared a feature in the changing dimension. For example, there could be two or three objects of the same shape or colour (shape and colour were both repeated, regardless of the change type). This encouraged participants to use similarity groupings to represent the changing dimension in VWM (e.g., Peterson & Berryhill, 2013). For example, participants could store representations of groups of identical colours in VWM to effectively detect colour changes, although the representation would not contain the shape information. Therefore, we predicted that the probability information would not be learned and used in Experiment 1, which would replicate the results of Beck et al. (2008).

Method

Participants. Across all experiments, some participants reported not having normal or corrected-to-normal vision, or did not report normal colour vision. These participants were excluded from analyses. In all experiments, we report how many participants were excluded due to non-normal vision, followed by the total number of participants who were included in the study after these participants were excluded. In Experiment 1, three participants were excluded due to non-normal vision, for a final total of 74 undergraduate students who participated in this experiment for credit in their undergraduate courses, (51 female, average age = 20 years). Thirty-seven students were randomly assigned to the colour predictive condition (17 in the red condition, 20 in the blue condition), and 37 randomly were assigned to the shape predictive condition (17 in the Shape 1 condition and 20 in the Shape 2 condition). All participants that are included self-reported normal or corrected to normal vision and normal colour vision.

Stimuli. Stimuli consisted of four shapes from the Fiser and Aslin (2002) set of shapes in four unique colours (red, green, blue, and yellow), for a total of 16 unique colour–shape combinations. For each experiment, 164 memory-display/test-display pairs were created: 120 for the predictable trials and 44 for the unpredictable trials. Each display contained six coloured shapes in a circle of 12 possible locations; the location of each object was selected at random. From a viewing distance of approximately 47 cm, each object was at a 3° visual angle and the diameter of the circle was at a 9.6° visual angle. Each unoccupied location was filled with a black square and participants were instructed to ignore the black squares and attend only to the coloured shapes. Empty spaces were filled with black squares to discourage configural groupings.

In each display, features were repeated in the changing dimension to encourage feature similarity groupings in that dimension; however, features were repeated in both the predictive and changing dimension. In the predictive dimension, all of the possible features (including the predictive feature) appeared at least once in the memory displays (all features in the predictive dimension, including the predictive feature, appeared on average 1.5 times across all displays). At least three out of the four features in the changing dimension appeared in the memory display. Therefore, in the colour predictive (shape change) condition, all four colours and at least three of the four shapes were present in the memory display. In the shape predictive condition, all four shapes and at least three of the four colours were present. Therefore, any given feature in the predictive dimension appeared between one and three times and any given feature in the changing dimension appeared between zero and four times in each memory and test display. Importantly, for all displays in both conditions, at least one colour and one shape appeared at least twice. These constraints were similar to those in Beck et al. (2008).

The test display for each memory-display/test-display pair was identical to its corresponding memory display with the exception that a single object changed to a randomly determined shape (colour predictive condition) or colour (shape predictive condition). On the predictable trials, the object that changed was selected at random from all objects in the memory display that contained predictive feature. On the unpredictable trials, the object that changed was selected at random from all objects in the memory display that did not contain the predictive feature. That is, if there were two or more objects that shared a feature in the predictive dimension (e.g., two red objects), the object that was selected to be the changing object was determined randomly (e.g., either red object was equally likely to change).

Procedure. Participants saw a memory display of six objects for 2,000 ms, followed by an 800-ms interstimulus interval (ISI), and then the test display for 2,000 ms (see Figure 1). On each trial, a single object either changed colour (shape predictive) or shape (colour predictive). Following the test display, participants viewed a response screen that contained placeholder squares in all 12 locations; the squares contained letters and symbols that corresponded to the top row of letters on their keyboard. Participants selected the location of the changing object by pressing the corresponding key.

Participants completed trials of either all colour changes or all shape changes. The dimension that changed for that participant was the *changing dimension*. The *predictive dimension* was always the nonchanging dimension. Therefore, for participants who completed colour change trials, the predictive dimension was shape, and for participants who completed shape change trials, the predictive dimension was colour. In each dimension, two features were selected to be predictive features, and participants were randomly assigned to one predictive feature condition. When detecting shape changes, colour was predictive, and either red or blue was the predictive colour. When detecting colour changes, shape was predictive, and either Shape 1 or Shape 2 (shapes arbitrarily labelled, see Figure 1 for an example of each of these shapes) was the predictive shape.

Participants first completed 32 practice trials, which contained 24 predictable trials and eight unpredictable trials, randomly intermixed. Practice trials were immediately followed by 120 high-weighted trials, split into 3 blocks (Blocks 1–3) of 40 trials (30 predictable trials and 10 unpredictable trials, presented in a random order). Following the high-weighted trials, participants completed one block of low-weighted trials (Block 4) of 12 trials (6 predictable trials and 6 unpredictable trials presented in a random order).

At the end of the experiment, participants completed a questionnaire to assess awareness of the probability information. The questions were presented one at a time on the computer screen and proceeded from general to specific to probe their awareness. The first question asked participants to "report any strategies you used

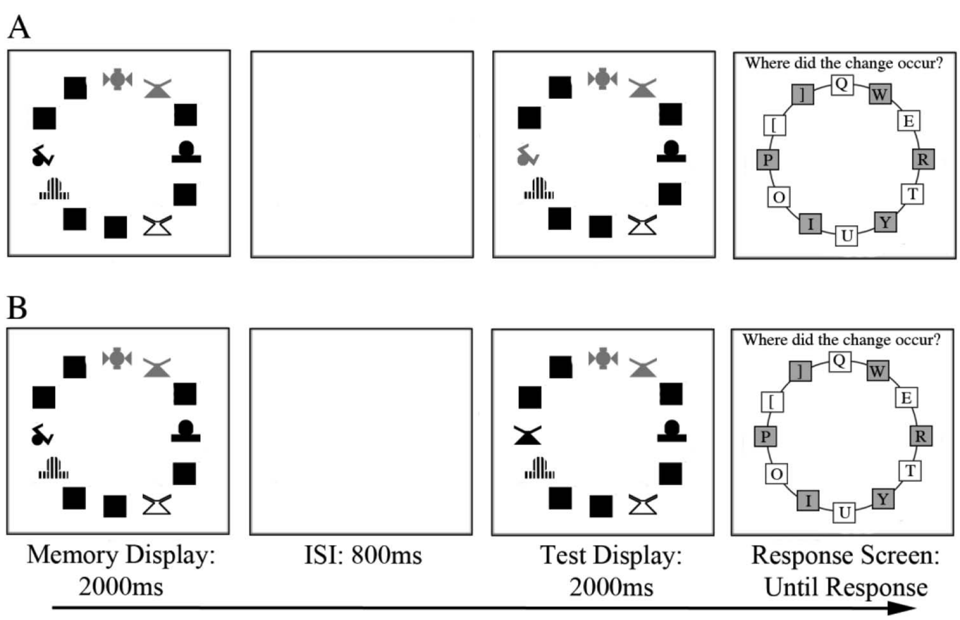

Figure 1. Procedure for Experiment 1. Figure 1A shows a colour change and Figure 1B shows a shape change. The object in the 12:00 position in both figures depicts the Shape 2 while the object in the 8:00 position in both figures depicts the Shape 1 shape. The changing object is located in the 9:00 position. Participants viewed the test display for 2,000 ms, which was followed by a response screen with placeholder boxes in the relative positions of the objects from the display. The placeholder boxes were filled with letters and/or symbols on their keyboard. Participants reported the location of the changing item by pressing the key on the keyboard that corresponded to that location. For example, the changing object in the 9:00 position contains the letter *P* so participants would press the letter *P* on their keyboard to correctly report the location of the change. The response keys began at the 12:00 position with the letter *Q* and proceeded clockwise on the response wheel and rightward on the keyboard, so that the response for the final 11:00 position was the "]" key.

to complete this experiment." The second question asked whether they believed some objects of some features (in the predictive dimension) changed more often than others. The wording for the colour predictive (shape predictive in parentheses) was as follows: "there were 4 objects that could have appeared in any of the arrays. When the objects changed shape (colour), do you think that any of the colours (shapes) changed shape (colour) more frequently than others?" If participants responded "no," they did not complete the rest of the questionnaire. If they responded "yes" to this second question, the third question asked the participant to choose which colour/shape changed most often (all possible features in the predictive dimension were displayed and participants completed a 4 alternative forced choice, 4 AFC, task). The final question asked participants to report the percentage of the trials for which they thought that the predictive feature changed. Participants were considered explicitly aware of the change information only if answered yes to the second question and the proceeded to correctly identify the predictive feature on the third question.

Results

To determine whether participants learned and used the probability information, we compared performance on the predictable trials to the unpredictable trials (see Figure 2). If participants were able to learn and use probability information to selectively encode the predictive objects, performance should be higher on the predictable trials than the unpredictable trials.

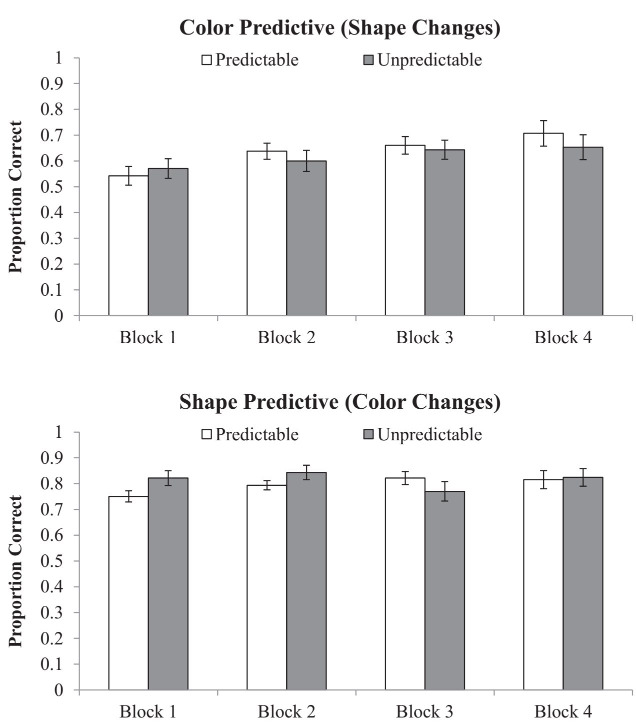

Figure 2. Results of Experiment 1. Blocks 1–3 were the high-weighted probability trials and Block 4 was the low-weighted probability trials. In Experiment 1, feature values were repeated. For both colour predictive and shape predictive, performance in the predictable trials was not higher than in the unpredictable trials. Error bars represent the standard error.

The same analyses were completed in all experiments. First, for each predictive dimension, we collapsed performance across the two predictive features (i.e., collapsed across red and blue when colour was predictive, and across Shape 1 and Shape 2 when shape was predictive[1]). Then, for each predictive dimension (colour or shape), we completed a 2 (probability) × 4 (block) within-subjects ANOVA with probability (predictable trials, unpredictable trials) and block (1–4) as within-subjects factors. We examined whether there was main effect of probability to determine if learning occurred. Any main effects of block that were found were the result of overall performance being higher in the later blocks than the earlier blocks, indicating that participants became better at the task overall with practice. We were also interested in whether there was a probability × block interaction, to determine how learning developed across time. For any instances in which the assumption of sphericity was violated, a Greenouse-Geisser correction was used, and epsilon values are reported.

In Experiment 1, probability information was not learned and used when colour was predictive and was only weakly and inconsistently learned and used when shape was predictive.[2]

Colour predictive (shape changes). For the colour predictive ANOVA, there was no main effect of probability, $F(1, 36) = 2.52$, $p = .12$, $\eta_p^2 = .07$. There was a main effect of block, $F(2.09, 75.18) = 6.97$, $p < .001$, $\eta_p^2 = .16$, $\varepsilon = .70$, as performance was lower in Block 1 than all other blocks (ps $< .01$). There was no probability × block interaction, which demonstrates that there was no change in the effect of probability across time, $F(3, 108) = 1.48$, $p = .23$, $\eta_p^2 = .04$.

Shape predictive (colour changes). There was no main effect of probability, $F(1, 36) = 2.29$, $p = .14$, $\eta_p^2 = .06$, and no main effect of block, $F(2.34, 84.21) = .89$, $p = .45$, $\eta_p^2 = .02$, $\varepsilon = .79$; however, there was a significant probability × block interaction, $F(3, 108) = 3.92$, $p = .01$, $\eta_p^2 = .10$. This was caused by higher performance on the unpredictable trials than on the predictable trials in Block 1 $t(36) = 2.71$, $p = .01$, and Block 2, $t(36) = 2.34$, $p = .02$, that was reversed in Block 3, $t(36) = 2.09$, $p = .04$, and was gone by the Block 4 (low-weighted block), $t(36) = .26$, $p = .79$.

Awareness questionnaire. In response to Question 1 on the questionnaire, only one participant (who was in the colour predictive condition) reported noticing the probability information and focusing on the predictable objects ("A lot of the time it was the blue shapes that changed. I paid greater attention to those once I realised this").

In response to the second question, 20 participants in the colour predictive (54%) and 10 participants in the shape predictive (27%) condition reported noticing that changes were more likely to happen to objects of some features more than others. Question 3 was used to determine whether the participants that answered yes

[1] In all experiments, an ANOVA was completed with probable feature as a between-subjects factor. Only in one instance did an ANVOA suggest that the probability effect was greater for one of the designated features (a significant probability × feature interaction that showed a probability effect in one feature greater than another). In Experiment 2, the probability effect appeared in Block 4 only for Shape 1 ($p = .17$ for Block 3), but the effect was significant in Blocks 2-4 for Shape 2.

[2] A separate ANOVA, with both predictability dimensions collapsed to maximize power, also failed to show a significant probability effect.

to Question 2 were then able to correctly identify the predictable feature. The predictive feature was accurately identified by 11 of the 20 (55%) participants in the colour predictive who answered Question 3, and six out of the 10 (60%) participants in the shape predictive condition. Therefore, 30% of all subjects in the colour predictive and 16% of all participants in shape predictive condition gave accurate responses to Questions 2 and 3.

To determine whether participants who answered Question 3 were more likely to select the predictable feature than would be expected by chance, a chi square goodness-of-fit test was completed for each condition, with the hypothesised rate of selecting the predictable feature 25% of the time, and an unpredictable feature 75% of the time. In both the colour predictive, X^2 (1, N = 20) = 9.60, $p < .01$, and shape predictive, X^2 (1, N = 10) = 6.53, p = .01 conditions, participants were more likely to report the predictable feature than would be expected by chance, suggesting that some participants were aware of this information.

Discussion

In Experiment 1, performance for the predictable trials was not consistently better than performance on the unpredictable trials. This indicates that probability information in the nonchanging (predictable) dimension was not incidentally learned and used to improve change detection performance; this was true for both colour and shape predictability. Although performance on predictable trials was higher than performance on the unpredictable trials for Block 3 when shape was the predictive dimension, the effect was not present on Blocks 1, 2, or 4 suggesting that if this is indicative of a probability effect, it was very slow to develop and was easily disrupted. Overall, these results replicate Beck et al. (2008), which suggest that learning and use of probability information does not readily occur if similarity groupings are encouraged in VWM. Although performance was generally good (~80% for colour change), it was not so high that a lack of a probability effect is likely to be attributable to ceiling effects: this is true especially for shape changes, where performance was much lower (~60%).

In response to the first question on the awareness questionnaire, some participants spontaneously reported using similarity groupings, for example "I tried to find the pairs or sets of the same shape" or "I would pair objects of the same colour." Many participants reported "grouping" (e.g., "grouping shapes together") or "finding patterns of colours," which may have been a reference to grouping based on similar features; however, the terms *grouping* and *pattern* are vague and could refer to other types of grouping strategies, such as spatial proximity. Therefore, we categorised responses as indicating grouping by feature similarity only when participants specifically reported used one or more of the following words: *identical*, *alike*, *similar*, *the same*, *pair/pairs*, *triplets*, *more than one*, *matched/matching*, *like*, *number of each*, *corresponding,* in addition to a reference to a feature or object (e.g., *similar colours*). Twenty-nine participants (39%) reported using such a similarity-grouping strategy (35% of the participants in the colour predictive condition and 43% of participants in the shape predictive), which offers support for the hypothesis that participants grouped together identical features in Experiment 1.

In Experiment 2, we tested the hypothesis that in the absence of similarity cues, participants would likely utilize connection groupings to represent information in VWM. This should allow them to encode the predictive dimension with the changing dimension (Luck & Vogel, 1997), which should facilitate learning of the probability information. Therefore, in Experiment 2, all of the features within the displays were unique (e.g., there could not be two red objects), to prevent similarity groupings. While it was expected that uniqueness in the changing dimension would be critical for encouraging connection groupings, features in both the predictive and changing dimensions were unique.

Experiment 2

The important difference between Experiment 2 and Experiment 1 was that all of the feature values within a display were unique in Experiment 2. There were a few other differences in the stimuli that are discussed below.

Method

Participants. Five participants were excluded due to nonnormal vision, for a final total of 92 undergraduate students (75 female, average age = 19.79 years) who participated in this experiment for credit in their undergraduate courses. Forty-six students were randomly assigned to the colour predictive condition of the experiment (24 in the blue condition and 22 in the red condition). Forty-six students were randomly assigned to the shape predictive condition (22 in the Shape 1 condition and 24 in the Shape 2 condition). All of the included participants self-reported normal or corrected to normal vision and normal colour vision.

Stimuli and procedure. The same stimuli and procedure were used as Experiment 1 with the following exceptions. Stimuli consisted of 12 shapes from the Fiser and Aslin (2002) set of shapes in 12 unique colours (red, dark green, lime green, burgundy, lavender, royal blue, turquoise, brown, grey, orange, pink, and yellow), for a total of 144 unique colour–shape combinations. There were no colour or shape repetitions within memory and test displays (see Figure 3).

As in Experiment 1, participants in the colour predictive (shape change) condition were randomly assigned to the blue or red conditions, and participants in the shape predictive condition were randomly assigned to the Shape 1 or Shape 2 conditions. In all conditions, the predictive feature (e.g., red in the red condition) only appeared in the predictable trials (24 of the 32 trials in the practice block, 90 of the 120 trials in the high-weighted probability blocks, and 6 of the 12 trials in the low-weighted probability blocks). On the unpredictable trials, the predictive feature did not appear at all, and any object could change. The postexperiment questionnaire was identical to that of Experiment 1, except that Question 3 was a 12 AFC task instead of a 4 AFC, of all possible features in the predictive dimension.

Results

The same two within-subjects ANOVAs were completed in Experiment 2 as in Experiment 1, and significant interactions were followed up with t tests comparing predictable and unpredictable trials in each block (see Figure 4). Probability information was learned and used, both when colour and shape were predictive.

Colour predictive (shape changes). The ANOVA revealed a main effect of probability, $F(1, 45) = 26.97$, $p < .001$, $\eta_p^2 = .38$:

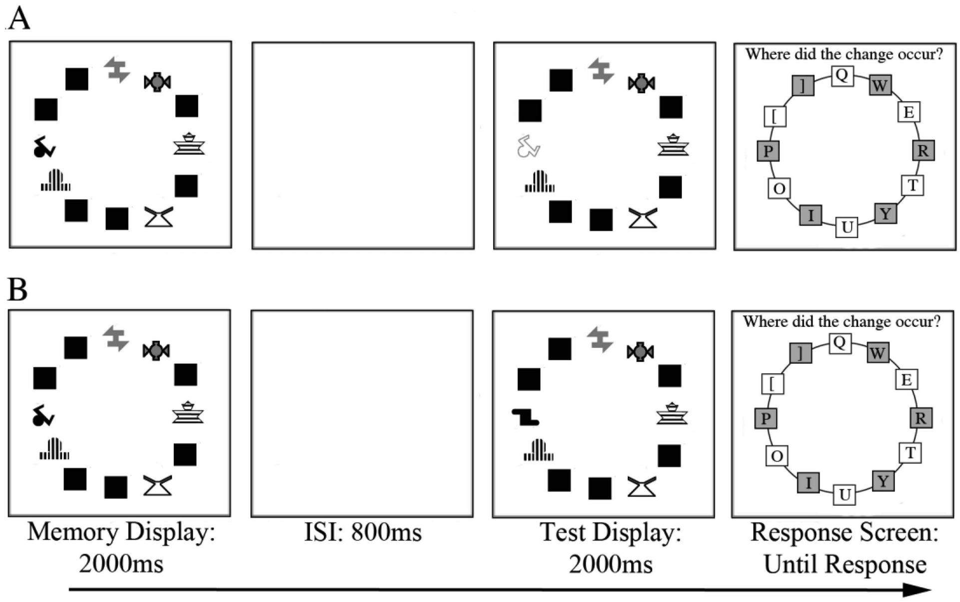

Figure 3. Procedure for Experiment 2. Figure 2A shows an example of a colour change and Figure 2B shows an example of a shape change. In this experiment, all feature values within a display were unique.

performance was higher for predictable trials than for unpredictable trials. There was no effect of block, $F(2.35, 105.93) = 1.62$, $p = .18$, $\eta_p^2 = .04$, $\varepsilon = .79$, and no interaction, $F(2.26, 101.77) = 1.36$, $p = .26$, $\eta_p^2 = .03$, $\varepsilon = .75$.

Shape predictive (colour changes). A main effect of probability was found, $F(1, 45) = 16.28$, $p < .001$, $\eta_p^2 = .27$: performance was higher for predictable trials than unpredictable trials. A main effect of block was also found, $F(2.32, 104.44) = 5.62$, $p = .001$, $\eta_p^2 = .11$, $\varepsilon = .77$. Performance was lower in Block 1 than all other blocks, *ps* $< .01$. Also, performance in Block 2 was lower than performance in Block 3, $p = .04$. Finally, a significant interaction was found, $F(2.52, 113.56) = 5.69$, $p = .001$, $\eta_p^2 = .11$, $\varepsilon = .84$, due to higher performance on predictable trials, but only in Blocks 3 and 4, both *ps* $< = .001$.

Awareness questionnaire. Sixteen participants (18%) reported noticing the probability information and using it to direct attention to the probable objects in response to Question 1 on the questionnaire: 10 participants (21%) in the colour predictive and six participants (13%) in the shape predictive condition. In response to Question 2, 36 participants (78%) in the colour predictive and 24 participants (52%) in the shape predictive responded that objects that contained certain features were more likely to change. Of these participants, 16 participants (67%) in the shape predictive and 33 participants (92%) in the colour predictive condition were able to correctly identify the predictive feature. Therefore, 78% of all subjects in the colour predictive and 52% of all subjects in the shape predictive conditions accurately responded to both Questions 2 and 3 to indicate explicit awareness of the probability information.

To determine whether participants who answered Question 3 were more likely to report the predictable feature, a chi square goodness-of-fit test was completed for each condition, with the hypothesised rate of selecting the predictable feature 8% of the time, and an unpredictable feature 92% of the time. In both the colour predictive, X^2 $(1, N = 36) = 327.27$, $p < .001$, and shape predictive, X^2 $(1, N = 24) = 106.91$, $p < .001$, conditions, participants were more likely to report the predictable feature than any nonpredictable feature, suggesting that those who reported noticing predictability information in Question 2 were aware of this information.

Discussion

In Experiment 2, performance was consistently higher for predictable trials, suggesting that probability information was learned and used to bias memory toward the predictive objects. In addition, the probability information continued to affect performance even after the probability information was reduced (Block 4). For the colour predictive (shape change) condition, there is a striking difference in the results between Experiment 1 and Experiment 2. There is a clear and consistent effect of probability in Experiment 2 and no effect in Experiment 1. In Experiment 2, in the shape predictive condition, a significant probability effect was found in Blocks 3 and 4, but not in Blocks 1 and 2, while in Experiment 1, there was a probability effect in Block 3. However, the size of the effect in Block 3 of Experiment 2 ($d = .70$) is larger than the effect in Block 3 of Experiment 1 ($d = .26$), and the effect persisted when the probability information changed in Block 4 for Experiment 2, but it did not persist for Experiment 1. Therefore, the probability effect for the shape predictive condition is stronger and more consistent in Experiment 2.

Comparison of performance in Experiments 1 and 2 (Experiment × Probability × Block mixed ANOVA) supports the prediction that similarity grouping was used in Experiment 1: overall performance was higher in Experiment 1 (when feature values were repeated) than in Experiment 2 (when all feature values were unique) for both the shape predictive, $F(1, 81) = 41.76$, $p < .001$, $\eta_p^2 = .19$, and colour predictive $F(1.81) = 19.15$, $p < .001$, $\eta_p^2 =$

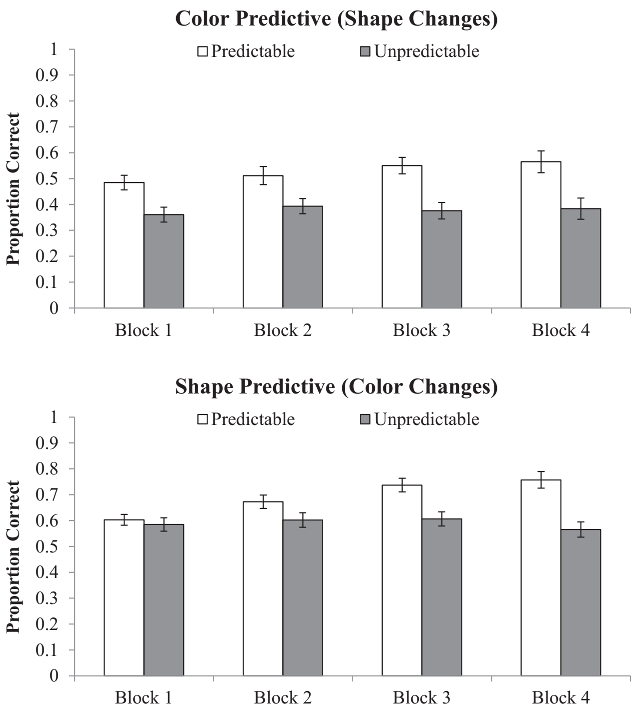

Figure 4. Results of Experiment 2. In Experiment 2, feature values were unique. For both colour predictive and shape predictive, performance was higher in the predictable trials than the unpredictable trials, although this reached significance only in Blocks 3 and 4. Error bars represent the standard error.

.19, conditions. This is supported by the responses to the first question in the questionnaire in Experiment 1, where many participants explicitly report using a similarity-grouping strategy.

However, another possible explanation for why learning occurred in Experiment 2, but not Experiment 1, is that the probability signal for the predictive feature was stronger in Experiment 2. That is, in Experiment 2, every time a predictive object was present, it changed. In contrast, Experiment 1, there were predictive objects present in the memory display that did not change. For example, in the case where red items were predictive in Experiment 1, there could be two red objects in the memory display, but only one of the red items changed shape at test. Furthermore, a red item could be present on an unpredictable trial. In contrast, in Experiment 2, only a single red item was present in the displays, and red only appeared in the memory displays on predictable trials; therefore, every time a red item was presented, it changed shape. Therefore, the probability of a predictive object changing during the high-weighted trials was 50% in Experiment 1 and 100% in Experiment 2. Thus, the probability information was more informative in Experiment 2 than in Experiment 1, because the probability signal was stronger in Experiment 2.

The question of whether the probability effect in Experiment 2 was caused by the type of VWM representation or the lack of noise in the probability signal was addressed in Experiments 3 and 4 (see Table 1). In Experiment 3, feature values were repeated in the memory display, as in Experiment 1, but all of the probable objects in the display changed in the predictive trials (e.g., all of the red objects within a display changed. There was still moderate noise in the probability signal (75% of predictive objects changed during the high-weighted trials), because the predictive object was present and did not change on unpredictable trials. In Experiment 4, unique features were used (as in Experiment 2), but while Experiment 2 had no noise in the probability signal, Experiment 4 had moderate noise in the probability signal: the predictive feature appeared, but did not change, in unpredictable trials (75% of the predictive objects changed during the high-weighted trials). Therefore, in both Experiments 3 and 4, 75% of all predictive objects changed during the high-weighted trials. If the strength of the probability signal is the reason that participants were unable to learn the probability information in Experiment 1, then learning should be the same in Experiments 3 and 4. However, if the representing the predictive feature in VWM is required for incidental learning of the probability information, then probability effects should only occur in Experiment 4.

Experiment 3

Experiment 3 was similar to Experiment 1: feature values within a display were repeated and the predictive feature was present on the unpredictable trials. However, in Experiment 3, multiple items could change from the memory to test display. Specifically, for a chosen feature (e.g., red), all objects that shared that feature would change. For example, in the red condition, if there were two red objects, both red objects would change on the predictable trials. On the unpredictable trials, any nonpredictable feature was randomly selected (e.g., green) and all objects with that feature would change

Table 1
Experimental Design Comparison of All Experiments

Experiment	Features	Percentage of predictive objects that changed	Predictive feature Present	Learning?
1	Repeat	50%	All trials	No
2	Unique	100%	Predictable trials	Yes
3	Repeat	75%	All trials	No
4	Unique	75%	All trials	Yes

Note. The "percentage of predictive objects that changed" column refers to the percentage of all predictable objects that appeared that would change, throughout the weighted probability blocks. This does not refer to the percentage of trials where the change occurred to an object with the predictable feature, which was always 75%. For example, in Experiment 1, if there were two predictable objects in a display, and only one of them changed, then 50% of the predictable objects on that trial changed.

(e.g., all of the green objects would change). Within any display, there could be one, two, or three changing objects. Participants were informed that they only had to localize one of the changing objects.

Method

Participants. Five participants were excluded due to non-normal vision, for a final total of 70 undergraduates (53 female, average age = 20.7 years) who participated in this experiment for credit in their undergraduate course. Thirty-four students were randomly assigned to the colour predictive (shape change) condition (20 in the blue condition and 14 in the red condition) and 36 were randomly assigned to the shape predictive (colour change) condition (18 in the Shape 1 condition and 18 in the Shape 2 condition) All participants self-reported normal or corrected-to-normal vision and normal colour vision.

Stimuli and procedure. The stimuli and procedure were identical to Experiment 1 with the following exceptions. Eight objects from a possible 16 (the same four colours and shapes as Experiment 1) were placed in a circle, with four black squares filling the empty locations. In each memory display, at least one value of each colour and shape was present, with the constraint that there could be no more than three of any colour or shape value (the predictive feature was repeated, on average, 1.96 times within a display). On a predictable trial, all objects with the predictive feature changed, and on an unpredictable trial, all objects sharing a randomly selected unpredictable feature changed. The feature that the object changed to was randomly selected from any of the possible features. When more than one object changed, participants were instructed that they only needed to report the location of one of the changing objects. On average, chance performance was 25% across predictable and unpredictable trials.

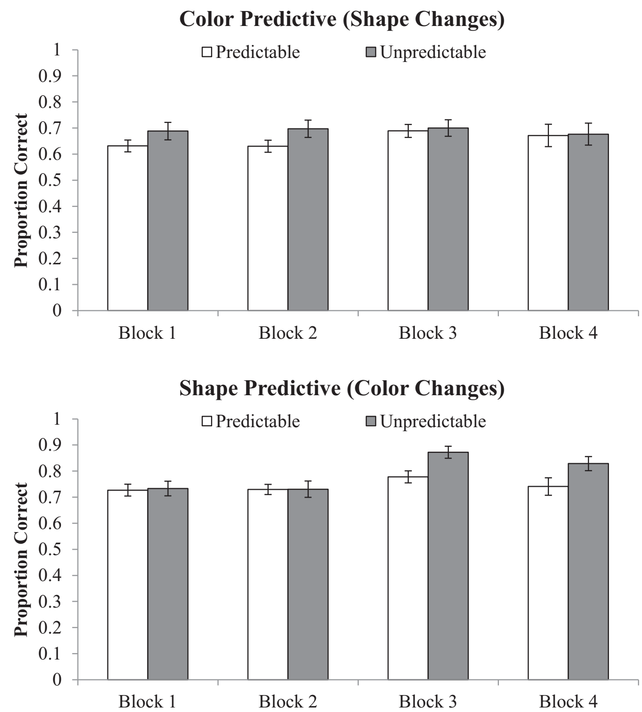

Figure 5. Results of Experiment 3. In Experiment 3, feature values were repeated. For both colour predictive and shape predictive, performance in the predictable trials was not overall higher than in the unpredictable trials. Error bars represent the standard error.

Results

Experiment 3 showed a similar pattern of results as Experiment 1: there was no evidence to support learning and use of probability information for either predictive dimension (see Figure 5).

Colour predictive (shape changes). There was a main effect of probability, $F(1, 33) = 4.35$, $p = .05$, $\eta_p^2 = .12$; however, this was caused by higher performance on unpredictable trials than on predictable trials. There was no main effect of block, $F(2.44, 80.39) = .66$, $p = .58$, $\eta_p^2 = .02$, $\varepsilon = .81$. There was also no interaction between block and probability, $F(1.99, 65.62) = 1.05$, $p = .37$ $\eta_p^2 = .03$, $\varepsilon = .66$.

Shape predictive (colour change). A main effect of probability was found, $F(1, 35) = 9.87$, $p = .003$, $\eta_p^2 = .22$. A main effect of block was also found, $F(1.86, 65.10) = 9.45$, $p < .001$, $\eta_p^2 = .21$, $\varepsilon = .62$: performance was lower in Block 1 and Block 2 than in Block 3 and Block 4. There was also an interaction between probability and block, $F(3, 105) = 2.72$, $p = .05$, $\eta_p^2 = .07$. Specifically, performance on the unpredictable trials was higher than performance on the predictable trials, but this was significant only in Block 3, $t(35) = 3.76$, $p = .001$ and Block 4, $t(35) = 2.08$, $p = .05$.

Explicit awareness. None of the participants reported noticing the probability information in response to Question 1. In response to Question 2, three participants (two in the colour predictive and one in the shape predictive), reported that objects that contained some features were more likely to change than others, but did not respond to Question 3 (or they responded, but entered a key that did not correspond to any of the choices); these participants were therefore excluded from subsequent analyses. Ten participants in the colour predictive condition (29%) and 17 participants in the shape predictive condition (47%) responded "yes" to Question 2, indicating that objects that contained certain features were more likely to change. Of these 27 participants, five participants (50%) in the colour predictive and two participants (12%) in the shape predictive correctly identified the predictive feature in response to Question 3. Therefore, 15% of all subjects in the colour predictive and 6% of all subjects in the shape predictive conditions accurately responded to both Questions 2 and 3 to indicate explicit awareness of the probability information.

To determine whether participants who answered Question 3 were more likely to select the predictable feature than an unpredictable feature, a chi square goodness-of-fit test was completed for each condition, with the hypothesised rate of selecting the predictable feature 25% of the time, and an unpredictable feature 75% of the time. In the colour predictive condition, participants were marginally more likely to select the predictable feature X^2 $(1, N = 10) = 3.33$, $p = .07$, although in the shape predictive condition, participants were equally likely to select the predictable or an unpredictable feature X^2 $(1, N = 17) = 1.59$, $p = .21$.

Discussion

Reducing the noise in the probability signal did not result in a change probability effect. This suggests that the results of Experiment 1 were not the result of a weak probability signal, and replicates the lack of a probability effect for displays with repeated features. Therefore, Experiment 3 further supports the conclusion that probability information is not incidentally learned and used when the changing dimension is stored in VWM at the exclusion of the predictive dimension. Furthermore, in Experiment 3, 19 participants (27%) reported using similarity-grouping strategies (32% in the colour predictive condition and 22% in the shape predictive condition), using the same criteria described in Experiment 1.

In some blocks, performance was higher on the unpredictable trials than the predictable trials. However, this effect was weaker ($\eta_p^2 = .12$ for colour predictive and $\eta_p^2 = .22$ for shape predictive) than the probability effect observed in Experiment 2 ($\eta_p^2 = .38$ for colour predictive and $\eta_p^2 = .27$ for shape predictive) and was inconsistent (Blocks 1 and 2 for colour predictive changes and Blocks 3 and 4 for shape predictive changes).

While participants who endorsed "yes" for Question 2 on the awareness questionnaire were marginally more likely to select the predictive feature in the colour predictive condition, this was not true in the shape predictive condition, suggesting a lack of awareness of the predictability information for these participants. The large number of participants reporting nonpredictive features could indicate that participants were attending to nonpredictive features on the predictable trials, lowering performance on the predictable trials.

In Experiment 3, the number of predictive objects that changed was 75%, which is more than in Experiment 1; however, in Experiment 2 the number of predictive objects that changed was 100%. This 100% predictability may be critical for learning to occur. Therefore, in Experiment 4, we once again used all unique features, but included the predictable feature in the unpredictable trials, so that in this case 75% of the predictable objects changed. Therefore, the strength of the probability information is the same in Experiments 3 and 4. If there is a probability effect in Experiment 4, the data will support the conclusion that the type of VWM representation is important for incidental learning of probability information.

Experiment 4

As in Experiment 2, all of the feature values within each array were unique, encouraging connection-grouping representations. However, unlike Experiment 2, a predictive object (e.g., a red object in the red condition) was present on the unpredictable trials, but was not the changing object.

Method

Participants. After exclusion of four participants due to nonnormal vision, a remaining 56 undergraduate students (37 female, average age = 19.84 years) participated in this experiment for credit in their psychology courses. Twenty-six participants were randomly assigned to the colour predictive condition (12 in the blue condition and 14 in the red condition). Thirty participants were randomly assigned to the shape predictive condition (16 in the Shape 1 condition and 14 in the Shape 2 condition). All participants self-reported normal or corrected to normal vision and normal colour vision.

Stimuli and procedure. The same stimuli and procedure were used as in Experiment 2 with the following changes. The predictive feature was present in all of the memory displays (e.g., red was always present in the red condition), but it would only change on the predictable trials (24 of the 32 trials in the practice block, 90 of the 120 trials in the high-weighted probability blocks, and 6 of the 12 trials in the low-weighted probability block). On the unpredictable trials, any randomly selected object that did not contain the predictive feature was the changing object.

Results

Replicating Experiment 2, we found evidence that probability information was learned and used to improve performance, for both predictive dimensions (see Figure 6).

Colour predictive (shape changes). There was a main effect of probability, $F(1, 25) = 18.32$, $p < .001$, $\eta^2 = .43$: performance was higher for predictable trials than unpredictable trials. No main effect of block was found, $F(2.25, 56.19) = .99$, $p = .40$, $\eta_p^2 = .04$, $\varepsilon = .75$, and there was no interaction $F(3, 75) = .24$, $p = .87$, $\eta_p^2 = .01$.

Shape predictive (colour changes). There was a marginal effect of probability, $F(1, 29) = 2.82$, $p = .10$, $\eta_p^2 = .09$; perfor-

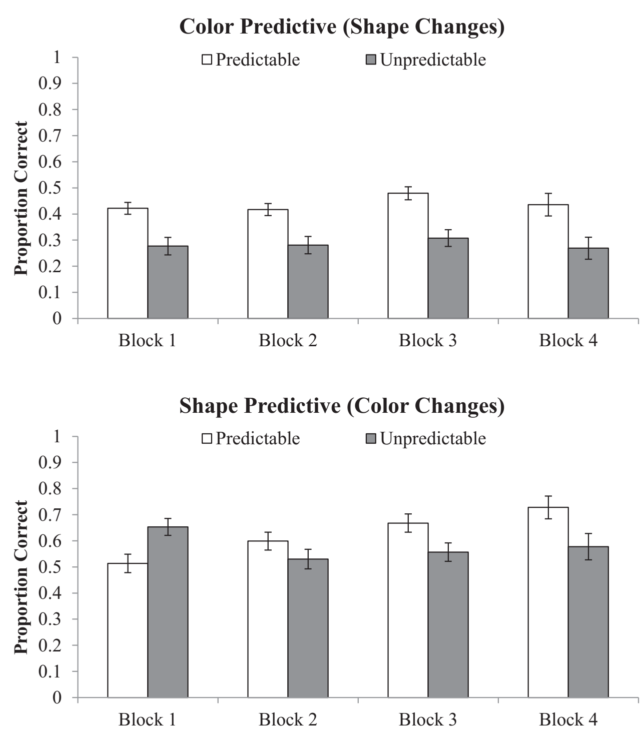

Figure 6. Results of Experiment 4. In Experiment 4, feature values were unique. For both colour predictive and shape predictive, performance was higher in the predictable trials than the unpredictable trials (for shape predictive this was significant in Blocks 3 and 4 only). Error bars represent the standard error.

mance was higher for predictable trials than unpredictable trials. A main effect of block was also found, $F(3, 87) = 4.22$, $p = .008$, $\eta_p^2 = .13$: performance was higher in Block 4 than Block 1, $p = .03$, and Block 2, $p = .007$. In addition, performance in Block 3 was higher than performance in Block 2, $p = .02$. Finally, an interaction was also found, $F(3, 87) = 10.63$, $p < .001$, $\eta_p^2 = .27$: as in Experiment 2, the probability effect was significant only in Blocks 3, $t(29) = 3.19$, $p = .003$, and 4, $t(29) = 2.40$, $p = .02$. The probability effect was not significant in Block 2, $t(29) = 1.65$, $p = .11$, and performance on the unpredictable trials was higher than performance in the predictable trials in Block 1, $t(29) = 4.32$, $p < .001$.

Explicit awareness. Eight participants (13%) reported noticing the probability information and using it to attend to the probable objects in response to Question 1 of the questionnaire. Six participants (22%) in the colour predictive and two participants (6%) in the shape predictive conditions reported this strategy.

In response to Question 2, six participants in the colour predictive condition reported "yes," but either did not respond to Question 3, or they responded but entered a key that did not correspond to any of the choices in Question 3, and were excluded from the following analyses. All six of these participants were in the "red" condition, and three of them did report noticing that red objects were more likely to change in response to Question 1. Therefore, it is likely that they would have correctly reported the red feature in response to Question 3. While these participants are not included in the following analyses, the results do not change if they are included. Thirteen participants (50%) in the colour predictive and 16 participants (53%) in the shape predictive responded that objects that contained certain features were more likely to change in response to Question 2. Of these participants who went on to answer Question 3, 13 participants (100%) in the colour predictive and seven participants (44%) in the shape predictive condition were able to correctly identify the predictive feature in Question 3. Therefore, 50% of all subjects in the colour predictive and 23% of all subjects in the shape predictive conditions accurately responded to both Questions 2 and 3 to indicate explicit awareness of the probability information.

To determine whether participants who answered Question 3 were more likely to select the predictable feature than an unpredictable feature, a chi square goodness-of-fit test was completed for each condition, with the hypothesised rate of selecting the predictable feature 8% of the time, and an unpredictable feature 92% of the time. In both the colour predictive, $X^2 = (1, N = 16) = 26.27$, $p < .001$, and shape predictive, $X^2\ (1, N = 13) = 143$, $p < .001$, conditions participants were more likely to select the predictable feature than would be expected by chance.

Discussion

Experiment 4 showed that it is not necessary for the predictable object to change 100% of the time in order for the probability information to be learned and used, provided feature values are unique. Furthermore, this experiment replicates Experiment 2, demonstrating the robustness of the ability to incidentally learn probability information in the predictive dimension (which is not explicitly task-relevant for the participant), and that this learning is maintained when the strength of the probability information is weakened.

General Discussion

Overall, the results show that probability information in the predictive dimension was learned and used to improve performance, but only when all features within a display were unique (Experiments 2 and 4). When features were repeated, probability information was not used (Experiments 1 and 3). This feature repetition factor may also be critical in understanding different results found across previous research: while in Beck et al. (2008), probability information about shape was not used to improve colour change detection (features were repeated), Droll et al. (2007) found that probability information about shape was used when items changed orientation (shapes were unique). Feature repetition is important because the type of VWM representation that is most efficient for the change detection task differs depending on whether or not features are repeated. Grouping by feature similarity (Peterson & Berryhill, 2013) is the most efficient VWM representation for detecting changes when there is repetition in the changing dimension. However, when there is no feature repetition, spatially connected groupings are efficient at representing all features of an object (Luck & Vogel, 1997), and this representation likely includes predictive dimension in VWM. Therefore, the ability to learn and use the probability information relies on utilizing a representational format in VWM that is most likely to include the predictive dimension.

We propose that participants did not use probability information in Experiments 1 and 3 due to a failure to encode the predictive dimension (at least in the same representation as the changing dimension). This may be a general attention effect that continues into VWM: participants may attend only to the changing dimension while the stimuli are present, which then leads to a failure to encode the predictive feature in the VWM representation. Therefore, repeating feature values may lead to a failure to attend and/or encode the predictive feature, precluding the learning and use of probability information.

Both the behavioural data and responses to the questionnaires support the hypothesis that participants used feature similarity-grouping representations in Experiments 1 and 3, while ignoring the predictive dimension. First, overall performance was higher in Experiment 1 than Experiment 2, suggesting that more of the changing features were stored in VWM in Experiment 1. Second, 39% of all participants in Experiment 1 and 27% of all participants in Experiment 3 explicitly reported using a similarity-grouping representation when asked what kinds of strategies they used to complete the task, although this does not necessarily mean that these were the only participants who engaged in similarity-grouping representations. Participants could have used similarity groupings without being explicitly aware that they were doing so, or without considering it a strategy, or may have been unable to adequately verbalize their strategy. Furthermore, Question 1 was very open-ended ("what strategies did you use") and we used a conservative criterion in determining that participants were reporting similarity groupings, so it is also possible that there were additional participants engaging similarity groupings that did not get coded as having used a similarity-grouping strategy. However, the relatively high rates of spontaneously reporting the use of similarity groupings at such an open question offers support for the hypothesis that participants engaged in grouping similar features to maximize capacity, while ignoring the predictive dimension.

The importance of feature uniqueness in learning and using probability information is unlikely to be caused by inconsistencies in the probability signal. Even when noise was held constant with 75% of the predictive objects changing, learning did not occur when features were repeated (Experiment 3) and learning did occur when features were unique (Experiment 4). It is therefore more likely that learning only occurs when VWM representations contain both the changing and predictive dimension.

The extent to which explicit awareness drives the learning effects is unclear. While awareness did occur for some participants, it is unclear whether knowledge arises from increased successful change detection following the use of probability information (e.g., "seeing" more changes after focusing on the predictive objects), or whether awareness comes first, leading participants to use it as a deliberate strategy. In this study, a questionnaire was introduced following the experiment that was designed only to determine whether any participants had gained awareness of the probability information. Future research could utilize more sensitive tests of awareness designed expressly to answer this question and larger groups of participants to separately analyse aware and unaware participants.

While Experiments 1 and 3 did not show the large, consistent probability effects shown in Experiments 2 and 4, there were some effects of probability. First, there were reversed probability effects, where performance was higher on predictable trials than unpredictable trials. It is possible that the reverse probability effects may have been driven in part by some subjects who believed that an unpredictable feature was predictable. In Experiment 3, some participants believed they noticed that they noticed probability information, but they were not any more likely to report the predictable feature than an unpredictable feature, suggesting that they were not aware of the probability information. It is possible that the participants who reported an unpredictable feature as most predictable biased VWM toward encoding objects of an unpredictable feature, thereby reducing performance on predictable trials. Second, there were some instances in which performance on predictable trials was higher than on unpredictable trials. These effects were numerically small and unlike Experiments 2 and 4, transient across blocks. It is therefore less clear whether these small and reversed effects might be theoretically meaningful or replicable.

The data presented here demonstrates that the ability to learn and use probability information is constrained by stimulus characteristics (to the extent that this influences the VWM representation). However, this may be particularly important for feature, rather than spatial, regularities. In Beck et al., (2008) information about the predictive location of a change was used even if the feature values of items were repeated. Furthermore, in van Lamsweerde and Beck (2011), predictive location information did not improve location change detection performance, possibly because location may be automatically included in the VWM representation. However, Olson, Jiang, and Moore (2005) found location change detection can improve when the same displays were repeated, suggesting a prioritization of the changing information. Therefore, it is possible that either location is automatically included in the VWM representation, or that learning spatial regularities does not depend on VWM resources. The finding that feature repetition suppresses incidental probability learning, found in the current study, may therefore be specific to nonspatial feature information.

The main design differences between the experiments where leaning did occur (2 and 4) and did not occur (1 and 3) is the repetition in the changing dimension, which we consider the most likely reason for the difference in results between these experiments. However, there are two additional differences between these studies, which could have influenced the results. The first of these is the total number of features in the displays. In Experiments 1 and 3, only four colours and four shapes were used, while in Experiments 2 and 4, 12 colours and shapes were used. Four colours and shapes were used to maintain close experiment parameters to Beck et al. (2008). It is possible that the presence of fewer features somehow contributed to the lack of learning in Experiments 1 and 3. The mechanism by which more features makes learning easier is unclear, and seems a less likely candidate for the probability effects found in Experiments 2 and 4. Although we cannot know for sure from the current data, we think it is unlikely that the number of features contributed to the effects found here.

Besides repetition in the changing dimension, features were also repeated in the predictive dimension in Experiments 1 and 3 (as in Beck et al., 2008). If repetition in the predictive dimension is a contributing factor, a possible explanation is that repeating features somehow made the predictive dimension easier to ignore, preventing learning. However, it seems likely based on previous research, which indicates that task-relevant information is more likely to be stored in VWM (Droll et al., 2005; van Lamsweerde & Beck, 2011), that the task-relevant (changing) dimension was responsible for the effect. Therefore, while we cannot exclude the possibility that the total number of features and repetition in the predictive dimension contributed to the probability effect, a more parsimonious explanation is that the VWM representation changed based on information (repetition or uniqueness of features) in the task-relevant dimension.

It is possible to take advantage of multiple types of grouping cues to maximize VWM capacity, such as connection (Alvarez & Cavanagh, 2004; Luck & Vogel, 1997; Logie et al., 2009; Luria, Sessa, Gotler, Jolicoeur, & Dell'Acqua, 2010; Luria & Vogel, 2011; Vogel et al., 2001), proximity (Woodman et al., 2003), or similarity Peterson & Berryhill, 2013). However, both cue availability and task demands may influence what kinds of strategies are used. In the current study, only a single dimension would change, so similarity grouping was the most efficient way to remember task-relevant information. However, if both colour and shape are relevant, similarity groupings may no longer be the most efficient way to store information. For example, Luck and Vogel (1997) found evidence that participants used connection groupings, even when there was repetition of features within a display. Therefore, it may be possible to flexibly switch between available grouping cues, using some while ignoring others. However, it is not yet clear whether some cues may dominate over others, or whether there may be extensive flexibility in choosing among different grouping cues based on top-down goals. The data here support the idea that there is some flexibility in using different grouping strategies in VWM.

In sum, probability information carried by a nonchanging dimension was used to improve change detection performance, but only when features were all unique. We suggest that repetition of feature values biases attention toward grouping similar features, which results in a lack of attention to connected features. Alternatively, unique features allow for attention to the connected feature, which in this case, carried the probability information. This suggests that although learning was incidental and there was no explicit requirement to attend to the nonchanging dimension (connected feature) in this task, this feature carried the probability information and had to be attended and stored in order for learning to occur. Therefore, while learning of probability information can occur without the intent to learn, it cannot occur without attending to the probability information.

Résumé

Dans le cadre de cette étude, nous avons tenté de déterminer si la capacité d'apprendre de l'information sur les probabilités est affectée par le type de représentation gardé dans la mémoire de travail visuelle. Dans les quatre expériences, les participants devaient détecter des changements au niveau de l'affichage de formes de couleur. Alors que les participants ont détecté des changements dans une (1) dimension (par ex., la couleur), une caractéristique, provenant d'une deuxième dimension non-changeante (par ex., la forme), pouvait prédire l'objet qui était le plus susceptible de changer. Dans les expériences 1 et 3, les items pouvaient être regroupés selon leur similarité dans la dimension changeante pour tous les items (par ex., les couleurs et les formes étaient répétées), alors que dans les expériences 2 et 4, les items ne pouvaient pas été regroupés selon leur similarité puisque toutes les caractéristiques étaient uniques. L'information sur les probabilités provenant de la dimension prévisionnelle a été apprise et utilisée pour améliorer le rendement, mais seulement quand toutes les caractéristiques affichées en un même temps étaient uniques (expériences 2 et 4). Lorsqu'il était possible de faire des regroupements par similarité des caractéristiques dans la dimension changeante (par ex. 2 objets bleus figurant dans un ensemble), les participants étaient incapables d'apprendre l'information sur les probabilités et de l'utiliser pour améliorer leur rendement (Expériences 1 et 3). Les résultats suggèrent que l'information sur les probabilités peut être apprise dans une dimension qui n'est pas explicitement liée à la tâche, mais seulement quand l'information sur les probabilités est représentée avec la dimension changeante dans la mémoire de travail visuelle.

Mots-clés : apprentissage incident, mémoire de travail visuelle, caractéristiques, objets, attention.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15,* 106–111. http://dx.doi.org/10.1111/j.0963-7214.2004.01502006.x

Anderson, D. E., Vogel, E. K., & Awh, E. (2013). Selection and storage of perceptual groups is constrained by a discrete resource in working memory. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 824–835. http://dx.doi.org/10.1037/a0030094

Baker, C. I., Olson, C. R., & Behrmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science, 15,* 460–466. http://dx.doi.org/10.1111/j.0956-7976.2004.00702.x

Beck, M. R., Angelone, B. L., & Levin, D. T. (2004). Knowledge about the probability of change affects change detection performance. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 778–791.

Beck, M. R., Angelone, B. L., Levin, D. T., Peterson, M. S., & Varakin, D. A. (2008). Implicit learning for probable changes in a visual change detection task. *Consciousness and Cognition, 17,* 1192–1208. http://dx.doi.org/10.1016/j.concog.2008.06.011

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General, 138,* 487–502. http://dx.doi.org/10.1037/a0016797

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36,* 28–71. http://dx.doi.org/10.1006/cogp.1998.0681

Droll, J. A., Gigone, K., & Hayhoe, M. M. (2007). Learning where to direct gaze during change detection. *Journal of Vision, 7*(14), article 6. http://dx.doi.org/10.1167/7.14.6

Droll, J. A., Hayhoe, M. M., Triesch, J., & Sullivan, B. T. (2005). Task demands control acquisition and storage of visual information. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 1416–1438. http://dx.doi.org/10.1037/0096-1523.31.6.1416

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 458–467. http://dx.doi.org/10.1037/0278-7393.28.3.458

Huang, L., & Pashler, H. (2007). A Boolean map theory of visual attention. *Psychological Review, 114,* 599–631. http://dx.doi.org/10.1037/0033-295X.114.3.599

Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning. *The Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology, 54,* 1105–1124. http://dx.doi.org/10.1080/713756001

Jiménez, L., & Méndez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 236–259. http://dx.doi.org/10.1037/0278-7393.25.1.236

Kondo, A., & Saiki, J. (2012). Feature-specific encoding flexibility in visual working memory. *PLoS ONE, 7*(12), e50962. http://dx.doi.org/10.1371/journal.pone.0050962

Logie, R. H., Brockmole, J. R., & Vandenbroucke, A. R. E. (2009). Bound feature combinations in visual short-term memory are fragile but influence long-term memory. *Visual Cognition, 17,* 160–179. http://dx.doi.org/10.1080/13506280802228411

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279–281. http://dx.doi.org/10.1038/36846

Luria, R., Sessa, P., Gotler, A., Jolicoeur, P., & Dell'Acqua, R. (2010). Visual short-term memory capacity for simple and complex objects. *Journal of Cognitive Neuroscience, 22,* 496–512. http://dx.doi.org/10.1162/jocn.2009.21214

Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia, 49,* 1632–1639. http://dx.doi.org/10.1016/j.neuropsychologia.2010.11.031

Manginelli, A. A., Geringswald, F., & Pollmann, S. (2012). Visual search facilitation in repeated displays depends on visuospatial working memory. *Experimental Psychology, 59,* 47–54. http://dx.doi.org/10.1027/1618-3169/a000125

Manginelli, A. A., Langer, N., Klose, D., & Pollmann, S. (2013). Contextual cueing under working memory load: Selective interference of visuospatial load with expression of learning. *Attention, Perception, & Psychophysics, 75,* 1103–1117. http://dx.doi.org/10.3758/s13414-013-0466-5

Olson, I. R., Jiang, Y., & Moore, K. S. (2005). Associative learning improves visual working memory performance. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 889–900. http://dx.doi.org/10.1037/0096-1523.31.5.889

Peterson, D. J. & Berryhill, M. E. (2013). The Gestalt principle of similarity benefits visual working memory. *Psychonomic Bulletin & Review, 20,* 1282–1289. http://dx.doi.org/10.3758/s13423-013-0460-x

Turk-Browne, N. B., Jungé, J., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134,* 552–564. http://dx.doi.org/10.1037/0096-3445.134.4.552

van Lamsweerde, A. E., & Beck, M. R. (2011). The change probability effect: Incidental learning, adaptability, and shared visual working memory resources. *Consciousness and Cognition, 20,* 1676–1689. http://dx.doi.org/10.1016/j.concog.2011.09.003

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature, 428,* 748–751. http://dx.doi.org/10.1038/nature02447

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 92–114. http://dx.doi.org/10.1037/0096-1523.27.1.92

Wegener, D., Ehn, F., Aurich, M. K., Galashan, F. O., & Kreiter, A. K. (2008). Feature-based attention and the suppression of non-relevant object features. *Vision Research, 48,* 2696–2707. http://dx.doi.org/10.1016/j.visres.2008.08.021

Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review, 10,* 80–87. http://dx.doi.org/10.3758/BF03196470

Woodman, G. F., & Vogel, E. K. (2008). Selective storage and maintenance of an object's features in visual working memory. *Psychonomic Bulletin & Review, 15,* 223–229. http://dx.doi.org/10.3758/PBR.15.1.223

Xu, Y. (2006). Understanding the object benefit in visual short-term memory: The roles of feature proximity and connectedness. *Perception & Psychophysics, 68,* 815–828. http://dx.doi.org/10.3758/BF03193704

Received April 7, 2014
Accepted March 21, 2015 ■