Short Research Article

How Do People Code Information in Working Memory When Items Share Features?

Gaën Plancher[1] and Robert L. Goldstone[2]

[1]Laboratoire d'Etude des Mécanismes Cognitifs, Université Lumière Lyon 2, Bron, France
[2]Indiana University, Bloomington, IN, USA

Abstract. A large literature suggests that the way we process information is influenced by the categories that we have learned. We examined whether, when we try to uniquely encode items in working memory, the information encoded depends on the other stimuli being simultaneously learned. Participants were required to memorize unknown aliens, presented one at the time, for immediate recognition of their features. Some aliens, called twins, were organized into pairs that shared every feature (nondiscriminative feature) except one (discriminative feature), while some other aliens, called hermits, did not share feature. We reasoned that if people develop unsupervised categories by creating a category for a pair of aliens, we should observe better feature identification performance for nondiscriminative features compared to hermit features, but not compared to discriminative features. On the contrary, if distinguishing features draw attention, we should observe better performance when a discriminative rather than nondiscriminative feature was probed. Overall, our results suggest that when items share features, people code items in working memory by focusing on similarities between items, establishing clusters of items in an unsupervised fashion not requiring feedback on cluster membership.

Keywords: working memory, category learning, acquired equivalence, memory

OPEN DATA OPEN MATERIALS

The human memory system can be described in terms of the flow of information that goes into and out of a short-term store, where the short-term store (also called working memory) is viewed as a temporary activation of some portion of information present in a long-term store (Atkinson & Shiffrin, 1971). We have known for decades that interactions between long-term memory and short-term memory are bidirectional. Retrieval from long-term memory benefits from processes occurring in working memory, such as attentional maintenance (Camos & Portrat, 2014; McCabe, 2008) and elaborative rehearsal (Gardiner, Gawlik, & Richardson-Klavern, 1994). Also, as a system with limited capacity, working memory takes advantage of processes that reduce the quantity of information needed to be maintained. For example, working memory performance is better when the materials that participants are required to maintain have already been stored in long-term memory (e.g., Loaiza, Duperreault, Rhodes, & McCabe, 2014), or when they can create higher-level chunks out of lower-level elements (Miller, 1956; Simon, 1974) because a chunk corresponds to a familiar pattern already stored in long-term memory (Ericsson & Kintsch, 1995). When items share features, they can be grouped into a small number of categories due to compression of information in working memory (Chekaf, Cowan, & Mathy, 2016; Thalmann, Souza, & Oberauer, 2019), thereby reducing the working memory load.

However, even if it is easier to actively maintain working memory information already stored in long-term memory, the purpose of the present study was to examine whether, when people encode isolated items, the information encoded would depend on the other stimuli being simultaneously learned. For example, if we try to remember that an orange just appeared, then a different feature (circular shape) should be remembered if the other items could have been (goldfish, carrot, pencil) than the feature (color) that should be remembered if the other items could have been (moon, baseball, clock). The working memory literature shows that, when presented in the same list, similar items (phonologically or visually) are more poorly memorized than dissimilar items (Baddeley, 1966; Logie, Della Sala, Wynn, & Baddeley, 2000). Also, items in a list belonging to the same region in a semantic network are prone to producing false recognitions (Atkins & Reuter-Lorenz, 2008). However, research on working memory has not paid sufficient attention to determinants of what is encoded. The objective of the present study is to investigate how people code items in working memory when items share features.

A large literature suggests that the way we process information is influenced by the categories that we have learned. The classic categorical perception effect shows that discrimination of stimuli that belong to different categories is improved after category learning training (i.e., acquired distinctiveness), while the ability to discriminate among stimuli belonging to the same category is reduced (i.e., acquired equivalence; Goldstone, 1994; Goldstone, Lippa, & Shiffrin, 2001; Harnad, 1987). The categorical perception effect is typically found using a supervised training procedure, where feedback about the category structure is provided to participants. However, categorical perception is also found when categories are not explicitly provided but rather indicated by the unsupervised organization of examples into separable clusters (Gureckis & Goldstone, 2008). That is, if items belonging to a single supervised category fall into two clusters according to their visual features, then discrimination is better when the items belong to different, rather than the same, clusters. Categorical perception can be viewed as a useful strategy because it reduces the amount of information that needs to be stored from a rich perceptual input. The perceptual information that distinguishes between different members of the same category can be discarded if these differences are not relevant for a subsequent categorization task. For example, the idiosyncratic variations of hand drawn letter "a"s may be relatively unimportant for word recognition and can be safely ignored.

This should be pertinent, in particular, for cognitive systems that rely on limited resources, such as working memory. One strategy for reducing the quantity of new information needed to be maintained in working memory could be to focus on similarities between items, establishing clusters of items in an unsupervised fashion not requiring feedback. In this case, the quantity of information to be coded would be reduced because only cluster-level information would be stored rather than the more elaborate and detailed encodings needed to uniquely identify items within a cluster. This compression would reinforce the learning of shared, nondiscriminative features – features that do not discriminate between items that belong to the same unsupervised category based on their similarities. However, if compressing items into categories decreases the ability to discriminate items within a category, we would also expect confusions among items belonging to the same category.

Another strategy would be to focus on how each item is distinctive from the others in its category. This strategy could still reduce the amount of to-be-maintained information by selectively drawing attention to discriminative rather than nondiscriminative features. In the differentiation models, information about new experiences with a stimulus in a particular context accumulates in a single memory trace and these updated memory traces become more distinct from the representations of other stimuli (Kılıç, Criss, Malmberg, & Shiffrin, 2017; Shiffrin & Steyvers, 1997). For example, the retrieving effectively from memory (REM) model (Shiffrin & Steyvers, 1997) assumes that the strength of a test item is based not only on the fact that it shares features with studied items but also on the diagnosticity (i.e., distinctiveness) of these matching features. A match of an uncommon, highly diagnostic feature provides higher evidence that the test item was studied than a match of a common, less diagnostic feature. Such hypotheses are supported by findings that words with uncommon letters are recognized better than words with common letters when word frequency is held constant probably because distinctive features attract more attention (Malmberg & Nelson, 2003). This kind of model can also explain the *strength-based mirror effect* that with an increasing number of presentations, both greater hit rates and lower false alarm rates are observed (Criss & Shiffrin, 2004). When learning to categorize stimuli, differentiation can operate by differentially weighting relevant feature dimensions (e.g., Nosofsky, 1987) or dimension differentiation, in which once integrated feature dimensions become more psychologically separable (e.g., Goldstone & Steyvers, 2001). The study of Kang and Pashler (2012) also stresses highlighting features that distinguish between different trained items. Through a comparison of spaced versus massed learning, their results suggested that spaced learning gives better performance as it helps participants during learning to focus on discriminative contrasts between stimuli.

The present research was motivated by two questions. First, are discriminative features more likely to be encoded in working memory than nondiscriminative features? Second, do feature values that discriminate between items become more or less discriminable in working memory with training? Participants were required to memorize unknown aliens, presented one at the time, for immediate feature recognition. Ten aliens in total were presented in the whole experiment. Most of the aliens were organized into pairs that shared every feature except one, which is the discriminative feature for the pair. Regarding the first question, if features that distinguish one object from similar alternative objects draw attention (the *distinctiveness weighting* hypothesis), we should observe better performance when a discriminative rather than nondiscriminative feature was probed. On the contrary, if people develop unsupervised categories (the *cluster encoding* hypothesis), creating a category for a pair of aliens, then we should observe better feature identification performance for nondiscriminative features because they are common to the two aliens belonging to the same pair. Regarding the second question, discriminating features will become increasingly encoded with training if people focus on them. However, discriminating values will become less discriminable with training if two values that belong to the same category items (pair) become more similar with training via acquired equivalence.

Method

Participants and Design

Eighty-two undergraduate psychology students (33 females, mean age = 19.30) from Indiana University participated for course credit. Four participants were excluded from further analyses because their average level of performance at the memory task was under 40% accuracy. The type of probe (discriminative vs. nondiscriminative) was manipulated within-subjects. The time interval was broken down into five blocks.

Materials

Two sets of ten aliens were created based on the material developed by Carvalho and Goldstone (2017). The two sets were created using the same features, and the difference between the two sets was the association between the features in a given alien. Half of the participants saw one set, and the other half saw the other set. Aliens had five different dimensions (arms, legs, eyes, mouth, belt, antenna). Each dimension had five different feature values possible. All aliens shared the same feature value for two dimensions: arms and legs. In other words, the arms and legs had the same appearance for all aliens. The ten aliens were split into five pairs. Four pairs contained *twin* aliens, and one pair contained *hermit* aliens. In the pairs of twins, the two aliens shared the same feature with their twin for three dimensions (e.g., asymmetrical eyes, smiling mouth, spiral belt) but had their own unique feature for one dimension (e.g., musical antenna vs. leaf antenna). In order to have every feature presented the same number of times, two hermits were created. They shared no features together (other than the features shared by all aliens) but shared each of their features with one other alien (see Figure 1A and B for some examples). The probes were generated in the same way for all alien types. That is, one of the four dimensions was randomly selected, and the five features associated with the selected dimension were presented.

Procedure

There were a study phase and a test phase: the alien was presented at study, and a probe at test. Trials began with a fixation dot centrally displayed on the screen for 500 ms.

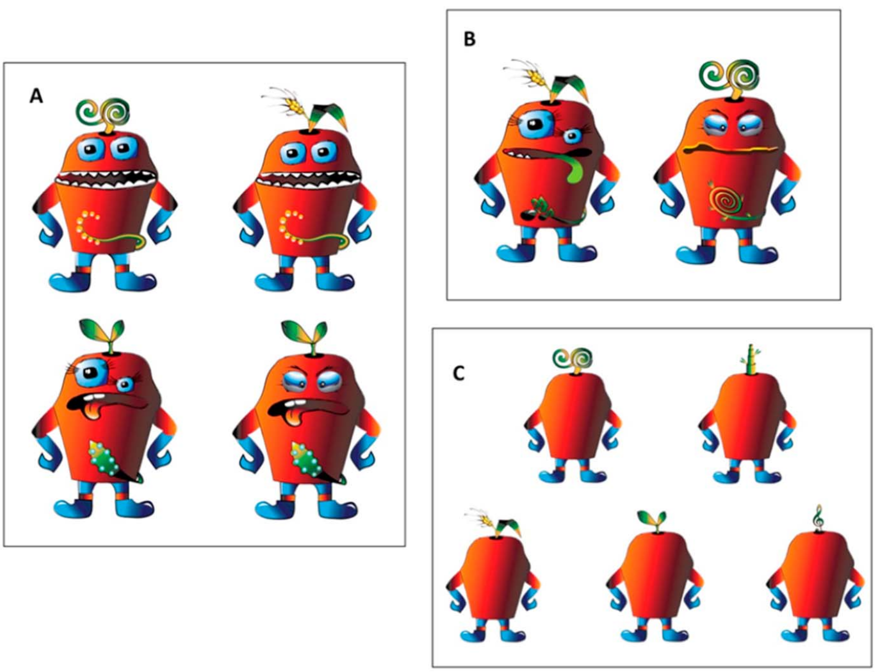

Figure 1. Two example pairs of twin aliens (A), one pair of hermit aliens (B), and a dimension probed at test, here *antenna* (C).

Then, a randomly selected alien was presented for 2,000 ms, followed by a 1,000 ms delay. At the end of a trial, one of the four dimensions was randomly selected, and the five features associated with the selected dimension were probed for a maximum of 8,000 ms (see Figure 1C for an example). On average, the participants provided a response in 1,700 ms, and as soon as they responded, the probe was removed from screen. Participants were instructed to remember the alien because they would later have to choose which feature belonged to the alien. They were required to be as accurate and fast as possible, but accuracy was presented as more important. At test, participants were instructed to click on the correct feature. A black rectangle appeared around the feature chosen. Each of the 10 aliens was presented 32 times in a random fashion. Every dimension was probed 80 times, also in a random fashion, and equally often for every alien. Before the 320 experimental trials, participants were invited to perform four practice trials. During these practice trials, different aliens were used. The aliens used for practice trials have the same body but different arms, legs, and other features than the ones used for experimental trials.

Accuracy was scored for *discriminative probe*, *nondiscriminative probe*, and *hermit probe*. The status of discriminative probe or nondiscriminative probe was relevant only for pairs of twins. A probe was considered as discriminative when it probed the dimension that was diagnostic for the current alien relative to its paired twin (e.g., the antenna for the aliens in the top row of Figure 1A; letters Q and R in Table 1). A probe was nondiscriminative when it probed a dimension that was nondiagnostic for the current alien (e.g., belt in Figure 1A; letters A, B, C, and D in Table 1). A probe was called a hermit probe when a hermit was presented on the trial. Different types of errors were also classified. If participants were probed for a feature shared with a twin, they could make errors when they wrongly chose the feature of the twin instead of the feature of the current alien (discriminative twin feature error) or when they chose another feature, which was not the one of the twin (discriminative nontwin feature error). They can also wrongly choose a feature for a nondiscriminative probe (nondiscriminative error). If participants were probed for a hermit, they could make an error by choosing any feature not present in the hermit (hermit error).

Table 1. Abstract design of the aliens.

Alien 1	AAAQ
Alien 2	AAAR
Alien 3	BBQB
Alien 4	BBRB
Alien 5	CQCC
Alien 6	CRCC
Alien 7	QDDD
Alien 8	RDDD
Alien 9	QQQQ
Alien 10	RRRR

Note. Every alien contains four dimensions (antenna, eyes, mouth, and belt); in the table, for every alien, one letter represents one dimension. Twins share the same feature for three dimensions (e.g., AAA for the first pair) but have their own unique feature for one dimension (i.e., Q or R). The first eight aliens have a twin. The last two aliens are hermits. Hermits (aliens 9 and 10) have no twin; each dimension for each hermit is shared with another alien to equate frequency across features. There is no relationship between the feature values along the different dimensions; letters such as A, B, and C are reused across the four columns just to make the pairs more obvious.

Results

First, analyses of variance on percent correct at the recognition task were conducted using the type of probe (discriminative, nondiscriminative, hermit) and blocks of trials (1, 2, 3, 4, 5) as within-subject variables (Figure 2). A main effect of the type of probe ($F(2, 140) = 3.27$, $p < .05$; $\eta^2 = 0.002$) showed better memory performance for discriminative and nondiscriminative probes compared to hermit probes. Individual contrasts showed that the difference between hermit and nondiscriminative probes was significant ($p < .05$), suggesting that having a twin, which is not the case for hermit, increased the likelihood of learning an alien well. But there was no difference between discriminative and hermit features ($p = .21$), and also, no difference between discriminative probe and nondiscriminative probe was observed ($p = .70$). A main effect of blocks of trials ($F(4, 280) = 12.36$, $p < .001$; $\eta^2 = 0.05$) showed a learning process across trials with a plateau

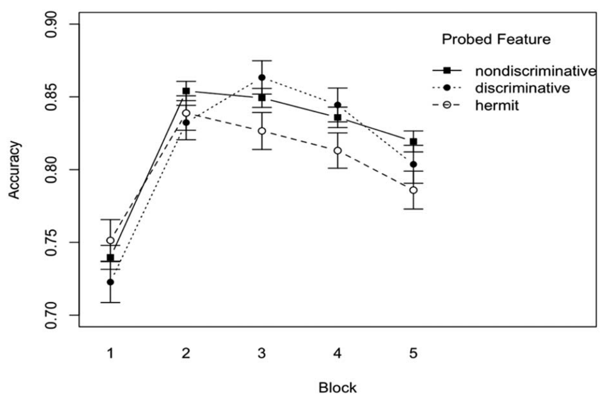

Figure 2. Accuracy at the recognition task according to the type of probe and the block (means and confidence intervals).

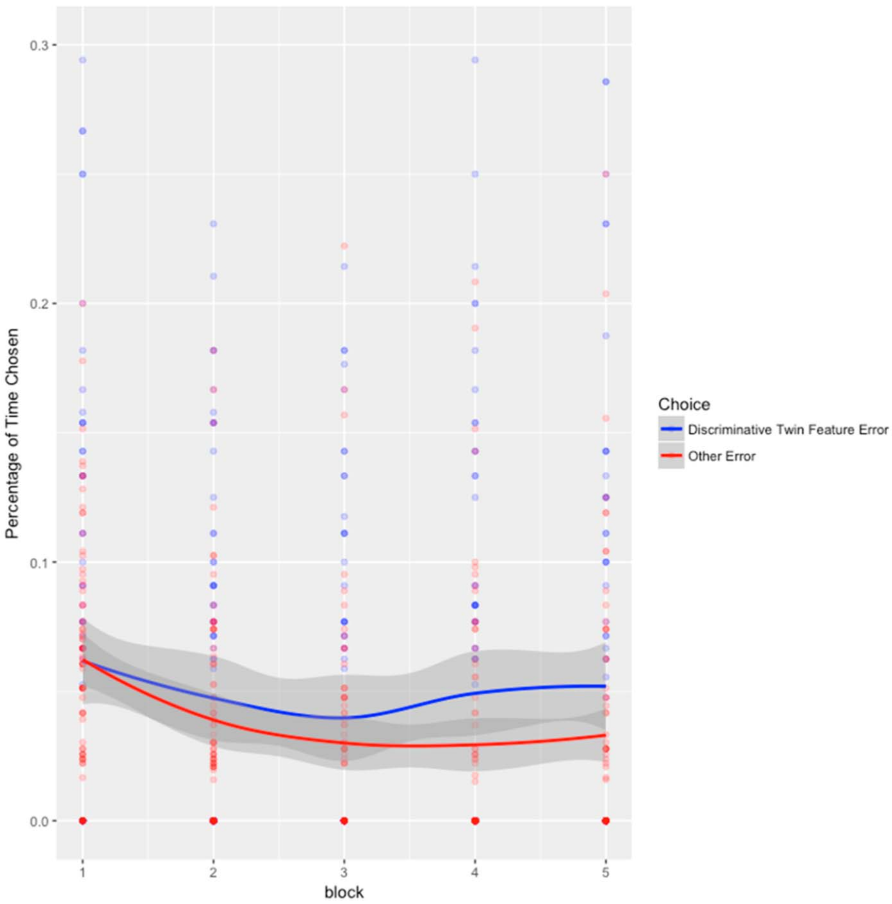

Figure 3. Proportion of errors across blocks of trials. The discriminative twin feature error was contrasted to the other types of errors (discriminative nontwin feature error, nondiscriminative error, hermit error).

occurring after the first block of trials. Individual contrasts showed that block 1 was significantly lower than the other ones (1 vs. 2, 3, and 4, $p < .0001$; 1 vs. 5, $p < .01$). The other blocks did not significantly differ among them, even if there was a tendency indicating better performance for block 3 compared to block 5 ($p = .06$). The interaction between the type of probe and blocks of trials was significant ($F(8, 560) = 1.99$, $p < .05$; $\eta^2 = 0.004$). We performed an exploratory ad hoc statistical test in which we grouped blocks 3, 4, and 5 and combined nondiscriminating and discriminating. We observed that combined discriminative and nondiscriminative features were better recalled than hermit features ($F(2, 148) = 6.84$, $p < .01$; $\eta^2 = 0.005$) after two blocks. This is consistent with the hypothesis that the commonly occurring features (discriminating and nondiscriminating) are better encoded than the rare (hermit) features.

Regarding the errors, we were particularly interested in trials where participants wrongly chose the twin part (discriminative twin feature error). This type of error may occur when a feature shared with a twin was probed. Accordingly, we contrasted the proportion of discriminative twin feature errors with an average of the other three errors (discriminative nontwin feature error, nondiscriminative error, hermit error). We normalized the results for the fact that, by chance, one would expect three times more other error choices than twin choices. Concretely, we divided the average of the other three errors by three. Analyses of variance on errors at the recognition task were conducted using the types of errors (twin part, other error) and blocks of trials (1, 2, 3, 4, 5) as a within-subject variables. A main effect of the type of error ($F(1, 70) = 17.02$, $p < .0001$; $\eta^2 = 0.05$) showed more errors involving choices of the twin part compared to other errors (see Figure 3). The effect of block ($F(1, 70) = 3.02$, $p = .09$; $\eta^2 = 0.02$) and the interaction between block and choice type was not significant ($F(1, 70) = 1.32$, $p = .25$; $\eta^2 = 0.007$).

Finally, to have more information about how participants learned the aliens, we tested whether participants were more correlated in their accuracy performance on aliens when two aliens came from the same pair (twin aliens) than if the aliens came from different pairs. We calculated the correlation over participants in the feature recognition accuracy for each pair of aliens. Early versus late correlations were compared. If participants tended to learn aliens of the same pair together, the within-pair correlations should be greater than the between-pair

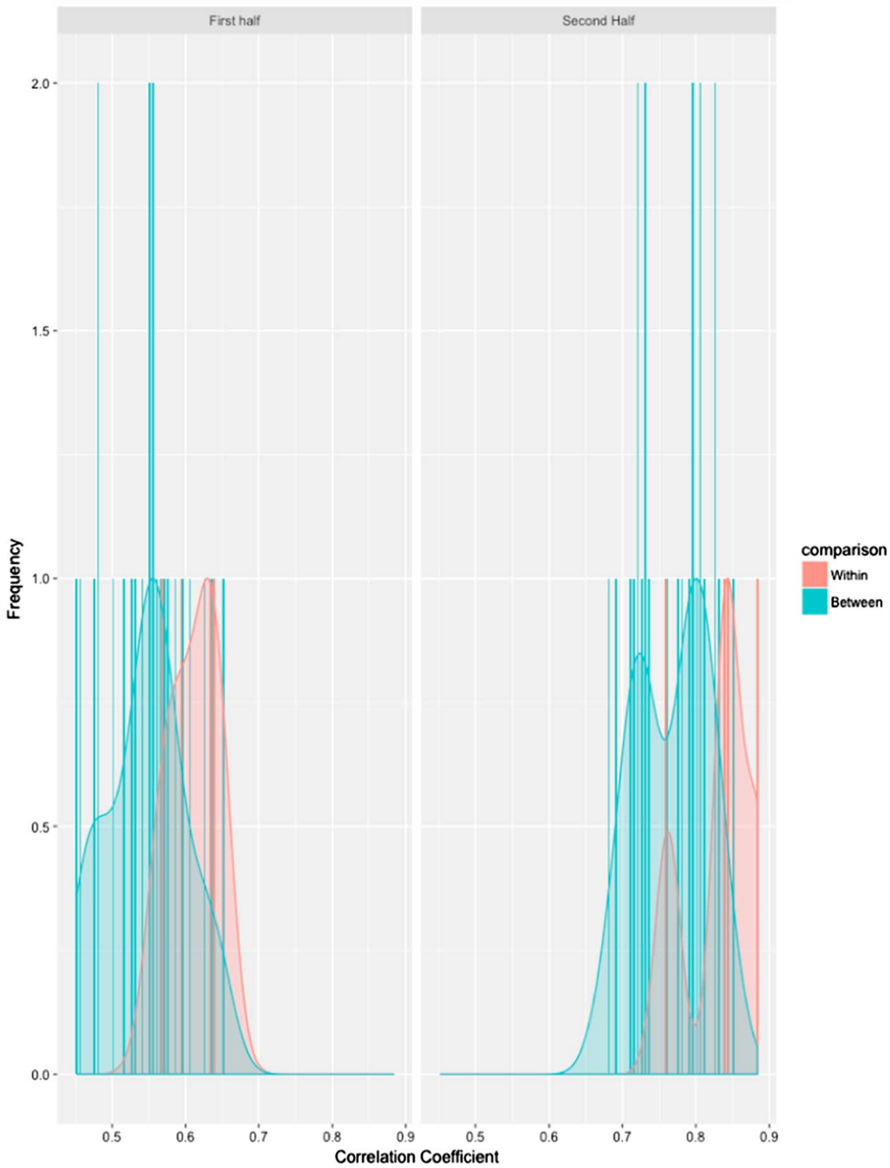

Figure 4. Distribution of correlation coefficients within pairs and between pairs of aliens for the first half and second half of the experiment.

correlations; also, the difference between within-pair correlations and between-pair correlations should increase over the trials (see Figure 4). The results showed that within-pair correlations were greater than between-pair correlations, and this was found for both halves. In particular, we observed that three within-pair correlations are found among the four largest correlations. Putting the hermits aside, there are 28 correlations: 4 within-pair correlations and 24 between-pair correlations. If the ranks of the correlations were distributed at random, with each rank (from 1 to 28, where 1 is assigned to the largest correlation and 28 to the weakest) being equally likely for each pair, then the probability of the event that we observe, namely, three within-pair correlations found among the four largest correlations, would be equal to:

$$\frac{1 + \binom{4}{3} \cdot 24}{\binom{28}{4}} \leq 0.005. \qquad (1)$$

In addition, there was some evidence that within-pair correlations were greater than between-pair correlations for the second half than the first half of trials.

Discussion

Because stimuli to be maintained in working memory are most of the time complex (e.g., a word contains information related to shape, sound, semantics, and phonology), it is likely that, to reduce working memory load, implicit choices are made while coding information for further maintenance. In this study, we wondered whether, when we try to encode isolated items, the information encoded depends on the other stimuli being simultaneously learned. In particular, we investigated how people code items in working memory when items share features. Participants were required to memorize unknown aliens, presented one at the time, for immediate recognition of their features. Most of the aliens were organized into pairs that shared every feature except one, which we call the discriminative feature for the pair. Overall, our results provided evidence in favor of the *cluster encoding* hypothesis, but not in favor of the *distinctiveness weighting* hypothesis.

Our findings suggest that attention was not focused only on features needed for isolated identification because accuracy for nondiscriminative features was equivalent to accuracy for discriminative features. It is likely that while participants encode items, they emphasize information shared by the two twins within a pair. Consistent with the idea that participants learn what is common between the twins, we found that after the first block, performance was greater for nondiscriminative probes compared to hermit, indicating that having a twin, which is not the case for hermit, increased the likelihood of effectively encoding the features that an alien shared with its twin. Also, people are apparently learning how to encode the aliens category-by-category to some degree because correlations in accuracy for parts in a twin were higher than correlations in accuracy for parts across twins. In addition, errors were greater for discriminative twin features than other kinds of probes. This difference was accentuated by training, suggesting that features within the same cluster (twin pair) acquired equivalence over time. Because we counterbalanced the stimuli so that one half of the participants saw one set of aliens and another half another set, and because each dimension (i.e., mouth, eyes, belt, antenna) is a discriminating feature for one set of twins, this result is probably not due to differences in salience for the discriminative parts.

All of these different findings are consistent with the idea that using unsupervised learning, participants code information to be maintained in working memory by focusing on what is common between items belonging to the same cluster rather than on what distinguishes between items belonging to the same cluster. These results point to learned category-level encodings of items for working memory, and not only for perception (Gureckis & Goldstone, 2008). The categories were not explicitly presented, but participants developed a strategy of trying to remember an item by creating an encoding for the twin-level category to which an item belonged. By picking up and using these unsupervised categories, participants created encodings that require less information, even though they are not maximally accurate, particularly for distinguishing between items belonging to the same category. Because nondiscriminative features are common to a category by definition, they are well encoded by subjects making category-level encodings. However, discriminations within a category were harder to make because arbitrarily grouped dimension values acquired equivalence over time (Goldstone, 1994), which explains why more errors were found for twin probes than other probes.

Our results are consistent with previous studies, suggesting that when items share features, people group them into categories and in so doing compress information in working memory (e.g., Chekaf et al., 2016; Thalmann et al., 2019). In the present case, people are apparently compressing information by creating categories even when items belonging to the same category were not presented together on a trial and no explicit feedback or highlighting of twin-level categories was given. As such, participants discover categories for themselves that can be used to efficiently code items. In the literature, two kinds of compressions have been distinguished. The first one refers to chunking of multiple, simultaneously present stimuli together to form a single code (e.g., Perruchet & Vinter, 1998). The second one refers to compressing multiple, sequentially presented stimuli because they come from the same cluster in a high-dimensional stimulus space (e.g., Mathy & Feldman, 2012; Chekaf, Cowan, & Mathy, 2016). The fact that in our study we can acquire equivalence effect between features that were never presented simultaneously can be predicted by the second but not first kind of compression.

Our results apply to working memory but also to learning, in general. Bates, Lerch, Sims, and Jacobs (2019) recently suggested that visual working memory is dynamically adaptive. They estimated peoples' visual working memory capacity by measuring their sensitivity to the distributions of stimulus features and their visual working memory sensitivity to the nature of a task such as which features are more or less important for performing the task. They used artificial plants as stimuli images, which varied in leaf widths and leaf angles. In one experiment, participants were first required to categorize plants, and one dimension was relevant (e.g., leaf width) for categorization and another was irrelevant (e.g., leaf angle). Participants were next required to perform a change detection task. They observed that participants' working memory was more sensitive to the dimension that was relevant to the categorization task relative to the task-irrelevant dimension. Their study also showed that the use of statistical regularities allows for more efficient memory because participants adapted their working memory performances to the specific distribution of dimension values. Visual working memory is thus shown to be task-dependent and dependent on learning. Persaud and Hemmer (2016) suggested that prior knowledge based on categories and uncertain working memory representations may be combined through Bayesian inference. Bayesian inference provides a normative theory for reasoning about and acting on uncertainty, but interestingly the rate-distortion theory goes further by providing a theory of processing when resources are limited (Bates et al., 2019). Bates et al. investigated how resources can be allocated in a manner that approaches the theoretical bounds on efficiency defined by information theory (e.g., Shannon & Weaver, 1949).

The assumption that working memory is adaptive is consistent with our finding that the general context of the experiment is taken into account when encoding an item on a single trial. When one item has to be isolated encoded, other stimuli to be learnt influence the item's maintenance. Creating categories to group stimuli is a more efficient strategy than learning items in an isolated way. Participants probably do not attempt to encode the information perfectly. That is why they also make errors, in particular confusing twin features because participants who create an efficient code to represent a twin-level category can theoretically respond with 100% accuracy to the nondiagnostic features shared by both members of a twin category, but only 50% accuracy for the diagnostic feature if they associate both diagnostic features with the twin category. This is also consistent with research showing that memory is sensitive to the statistics of visual information (e.g., Orhan & Jacobs, 2013).

To summarize, our study showed four main findings. First, twins were better recognized than hermits; second, this occurred because of learning as the first result did not appear on the first block; third, arbitrarily paired twin features with a dimension were more likely to be confused with each other than features not found in the same twin pair; and finally, twin seemed to be learned at about the same time – more so than expected by chance. Overall, this suggests that when items share features, people code items in working memory by focusing on similarities between items, establishing clusters of items in an unsupervised fashion not requiring feedback on cluster membership. This cluster-based categorization reinforces the learning of shared nondiscriminative features but also increases confusions among items belonging to the same category. This is consistent with the view that working memory is dynamically adaptive. In the future, it would be interesting to subsequently present the items to investigate whether, within a trial where items share similarities, they may be compressed to leave encoding resources for other items. For example, Brady, Konkle, and Alvarez (2009) observed that, within a trial, regularities among stimuli introduce redundancies that make the input more compressible, thus improving the encoding of other simultaneously presented items.

References

Atkins, A. S., & Reuter-Lorenz, P. A. (2008). False working memories? Semantic distortion in a mere 4 seconds. *Memory & Cognition*, *36*, 74–81. https://doi.org/10.3758/mc.36.1.74

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, *225*, 82–90. https://doi.org/10.1038/scientificamerican0871-82

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology*, *18*, 362–365. https://doi.org/10.1080/14640746608400055

Bates, C. J., Lerch, R. A., Sims, C. R., & Jacobs, R. A. (2019). Adaptive allocation of human visual working memory capacity during statistical and categorical learning. *Journal of Vision*, *19*, 11. https://doi.org/10.1167/19.2.11

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. https://doi.org/10.1037/a0016797

Camos, V., & Portrat, S. (2014). The impact of cognitive load on delayed recall, *Psychonomic Bulletin & Review*, *22*, 1029–1034. https://doi.org/10.3758/s13423-014-0772-5

Carvalho, P. F., & Goldstone, R. L. (2017). The sequence of study changes what information is attended to, encoded, and remembered during category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 1699–1719. https://doi.org/10.1037/xlm0000406

Chekaf, M., Cowan, N., & Mathy, F. (2016). Chunk formation in immediate memory and how it relates to data compression. *Cognition*, *155*, 96–107. https://doi.org/10.1016/j.cognition.2016.05.024

Criss, A. H., & Shiffrin, R. M. (2004). Interactions between study task, study time, and the low-frequency hit rate advantage in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 778–786. https://doi.org/10.1037/0278-7393.30.4.778

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211–245. https://doi.org/10.1037/0033-295x.102.2.211

Gardiner, J. M., Gawlik, B., & Richardson-Klavern, A. (1994). Maintenance rehearsal affects knowing, not remembering; elaborative rehearsal affects remembering, not knowing. *Psychonomic Bulleting & Review*, *1*, 107–110. https://doi.org/10.3758/bf03200764

Goldstone, R. (1994). Influences of categories on perceptual discrimination. *Journal of Experimental Psychology: General*, *123*, 178–200. https://doi.org/10.1037/0096-3445.123.2.178

Goldstone, R. L., Lippa, Y., & Shiffrin, R. M. (2001). Altering object representations through category learning. *Cognition*, *78*, 27–43. https://doi.org/10.1016/s0010-0277(00)00099-8

Goldstone, R. L., & Steyvers, M. (2001). The sensitization and differentiation of dimensions during category learning. *Journal of Experimental Psychology: General*, *130*, 116–139. https://doi.org/10.1037/0096-3445.130.1.116

Gureckis, T. M., & Goldstone, R. L. (2008). The effect of the internal structure of categories on perception. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 1876–1881). Austin, TX: Cognitive Science Society.

Harnad, S. (1987). *Categorical perception*. Cambridge, UK: Cambridge University Press.

Kılıç, A., Criss, A. H., Malmberg, K. J., & Shiffrin, R. M. (2017). Models that allow us to perceive the world more accurately also allow us to remember past events more accurately via differentiation. *Cognitive Psychology*, *92*, 65–86. https://doi.org/10.1016/j.cogpsych.2016.11.005

Kang, S. H. K., & Pashler, H. (2012). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology*, *26*, 97–103. https://doi.org/10.1002/acp.1801

Loaiza, V. M., Duperreault, K. A., Rhodes, M. G., & McCabe, D. P. (2014). Long-term semantic representations moderate the effect of attentional refreshing on episodic memory. *Psychonomic Bulletin & Review*, *22*, 274–280. https://doi.org/10.3758/s13423-014-0673-7

Logie, R.H., Della Sala, S., Wynn, V., & Baddeley, A. D. (2000). Visual similarity effects in immediate verbal serial recall. *Quarterly Journal of Experimental Psychology*, *53A*, 626–646. https://doi.org/10.1080/027249800410463

Malmberg, K. J., & Nelson, T. O. (2003). The word-frequency effect for recognition memory and the elevated-attention hypothesis. *Memory & Cognition*, *31*, 35–43. https://doi.org/10.3758/bf03196080

Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition*, *122*, 346–362. https://doi.org/10.1016/j.cognition.2011.11.003

McCabe, D. P. (2008). The role of covert retrieval in working memory span tasks: Evidence from delayed recall tests. *Journal of Memory and Language*, *58*, 480–494. https://doi.org/10.1016/j.jml.2007.04.004

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *101*, 343–352. https://doi.org/10.1037/h0043158

Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 87–108. https://doi.org/10.1037/0278-7393.13.1.87

Orhan, A. E., & Jacobs, R. A. (2013). A probabilistic clustering theory of the organization of visual short-term memory. *Psychological Review*, *120*, 297. https://doi.org/10.1037/a0031541

Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of Memory and Language*, *39*, 246–263. https://doi.org/10.1006/jmla.1998.2576

Persaud, K., & Hemmer, P. (2016). The dynamics of fidelity over the time course of long-term memory. *Cognitive Psychology*, *88*, 1–21. https://doi.org/10.1016/j.cogpsych.2016.05.003

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM-retrieving effectively from memory. *Psychonomic Bulletin & Review*, 145–166. https://doi.org/10.3758/bf03209391

Simon, H. A. (1974). How big is a chunk? By combining data from several experiments, a basic human memory unit can be identified and measured *Science*, *183*, 482–488. https://doi.org/10.1126/science.183.4124.482

Thalmann, M., Souza, A. S., & Oberauer, K. (2019). How does chunking help working memory, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*, 37–55. https://doi.org/10.1037/xlm0000578

History
Received October 2, 2019
Revision received April 1, 2020
Accepted April 6, 2020
Published online June 19, 2020

Open Data
The material used for the experimental procedure and the raw data of the present article are available on the following Open-Science Framework (OSF) web page: https://osf.io/j3dgf/.

Funding
This work has been supported by ANR-11-IDEX-0007, ANR-10-LABX-0042, to Gaën Plancher.

ORCID
Gaën Plancher
https://orcid.org/0000-0002-0178-6207

Gaën Plancher
Laboratoire d'Etude des Mécanismes Cognitifs
Université Lumière Lyon 2
5 avenue Pierre Mendès
69676 Bron
France
gaen.plancher@univ-lyon2.fr