# Benefits and pitfalls of data compression in visual working memory

Laura Lazartigues<sup>1</sup> · Frédéric Lavigne<sup>1</sup> · Carlos Aguilar<sup>1</sup> · Nelson Cowan<sup>2</sup> · Fabien Mathy<sup>1</sup>

Accepted: 14 May 2021 / Published online: 15 June 2021  $\odot$  The Psychonomic Society, Inc. 2021

#### Abstract

Data compression in memory is a cognitive process allowing participants to cope with complexity to reduce information load. However, previous studies have not yet considered the hypothesis that this process could also lead to over-simplifying information due to haphazard amplification of the compression process itself. For instance, we could expect that the over-regularized features of a visual scene could produce false recognition of patterns, not because of storage capacity limits but because of an errant compression process. To prompt memory compression in our participants, we used multielement visual displays for which the underlying information varied in compressibility. The compressibility of our material could vary depending on the number of common features between the multi-dimensional objects in the displays. We measured both accuracy and response times by probing memory representations with probes that we hypothesized could modify the participants' representations. We confirm that more compressible information facilitates performance, but a more novel finding is that compression can produce both typical memory errors and lengthened response times. Our findings provide clearer evidence of the forms of compression that participants carry out.

Keywords Visual · Working memory · Compression of information · Capacity · Response times

# Introduction

Working memory is known to be limited in both the amount of detail that can be retained in a visual scene (Bays et al., 2009; Brady & Alvarez, 2015; Ma et al., 2014; Schurgin et al., 2020) and the total number of objects that can be recalled regardless of how these objects can be detailed (Cowan, 2001; Luck & Vogel, 1997; Zhang & Luck, 2008). These two types of limitation are generally predicted by competing models such as those based on shared-resource or discreteslots, respectively (Bays & Husain, 2008; Rouder et al., 2008). However, the memory benefit caused by shared features in visual working memory does not seem to be accounted for easily by these two main classes of working memory models (Quinlan & Cohen, 2012), and other studies have shown that both types of limitations should be accounted

Laura Lazartigues Laura.lazartigues@univ-cotedazur.fr

for concurrently to fit data (Awh et al., 2007; Cowan et al., 2013; Hardman & Cowan, 2015; Oberauer & Eichenberger, 2013; Xu & Chun, 2006). An approach taken in the present study is to consider that models of capacity limits should better account for how shared features are processed in working memory, in particular when those features allow some room to be preserved in memory.

There are a few known factors that can help individuals recode information in a more efficient way, in particular when the visual scenes offer the possibility of associating features with one another during the task at hand (Gao et al., 2016; Jiang et al., 2000, 2004; Peterson & Berryhill, 2013; Woodman et al., 2003). For instance, the phenomenon of binding has been shown to allow objects to be integrated based on multiple features (Alvarez & Cavanagh, 2004; Bays et al., 2011; Fougnie & Alvarez, 2011; Saiki, 2019; Wheeler & Treisman, 2002; Xu, 2002). In the case of binding, let's imagine that one participant is presented with an array of four novel objects each made of three features, thus totaling 12 different features. Being able to recognize one of the multifeature objects does not mean that the participant would be able to recall precisely the three constituent features of each object. With a working memory capacity of four items, for instance, one could encode just two features of two different objects, or any other combinations such as three features for one object and only one feature for another object, but always



<sup>&</sup>lt;sup>1</sup> Department of Psychology, Université Côte d'Azur, BCL UMR 7320, CNRS, Campus Saint Jean d'Angely, MSHS, 24 avenue des diables bleus, 06357 Nice, Cedex 4, France

<sup>&</sup>lt;sup>2</sup> Department of Psychological Sciences, University of Missouri-Columbia, Columbia, MO, USA

totaling four features. In this case, a capacity of four items does not allow a given individual to retain the 12 features of the whole scene. However, multi-feature objects can be better memorized when features occur repeatedly over time and can be recoded as chunks. Once formed in long-term memory, those chunks can allow participants to hold a greater amount of information in working memory (Brady et al., 2009; Ngiam et al., 2019; Orbán et al., 2008). Provided that each multi-feature object chunk counts as one item, a participant with a capacity limit of four items could therefore perfectly recall an array of four recognized objects, each made of three features, thus recalling 12 different features.

In other cases, objects are not necessarily already encoded in long-term memory. For instance, previous studies have shown that bottom-up stimulus characteristics (e.g., Gestalt cues) can help participants group information to better recall stimulus items in visual scenes (Woodman et al., 2003). This is for instance the case with spatial information (De Lillo, 2004; Dry et al., 2012; Feldman, 1999; Haladjian & Mathy, 2015; Korjoukov et al., 2012; Sargent et al., 2010). Capacity in working memory can also be easily exceeded by perceptual organization when items share the same colors in a visual scene (Brady & Tenenbaum, 2013). Morey et al. (2015), for instance, showed that there is an advantage for singletons in displays containing repetitions (compared to displays containing no repetitions). Their interpretation was that grouping of repeated colors on the spot can leave room in working memory. One explanation of how grouping or chunking functions is that compression of information could be at work whenever room is being preserved in memory, particularly when individuals can find a way to efficiently recode information (Brady et al., 2009).

## The compression account

Compression of information involves recoding information into a more compact form. Although there is a controversy regarding whether visual working memory capacity is fixed regardless of information content (Alvarez & Cavanagh, 2004; Awh et al., 2007), the compression account predicts that storage is particularly efficient when information contains regularities. During the last decade, compression has been put forward to account for intelligence (Chekaf et al., 2018; Hutter, 2004), memory (Brady et al., 2009; Chekaf et al., 2016; Mathy & Feldman, 2012), language (Christiansen & Chater, 2016; Ferrer-i-Cancho et al., 2013; Kirby et al., 2015), and perception (Haladjian & Mathy, 2015; Nassar et al., 2018; Ramzaoui & Mathy, 2021). Compression is an information-theoretic concept based on algorithmic complexity. Algorithmic complexity corresponds to the shortest possible representation of an object (Li &

Vitányi, 2008). The gain offered by the shortest representation of an object (in comparison to the length of the original object) allows one to estimate the compressibility of the given object. Regarding the computational aspects of the theory, the shortest representation usually takes the form of the shortest program in a Turing machine. Although it is not computable (because one can never know whether maximal compression has been reached by a given recoding process), estimates can be obtained, see http://www.complexitycalculator (Gauvrit et al., 2016; Soler-Toscano et al., 2014). However, because different languages can be used interchangeably in place of Turing machines, indirect approaches can be taken by developing more practical metrics in a given domain. For instance, a compressibility metric has been developed in categorization to describe multi-feature objects (Feldman, 2003), and this metric (or very similar ones, see Kemp, 2012; Vigo, 2006) have proven useful to account for subjective complexity during learning (Bradmetz & Mathy, 2008; Feldman, 2000; Lafond et al. , 2007; Mathy & Bradmetz, 2004), so we decided to use this practical implementation in the present study to build our visual displays. The next section presents the metric and introduces the idea that a compression process can include pitfalls as well as benefits.

# The good and bad aspects of compression in working memory

To better understand how compression functions in working memory more globally, both advantages (optimization of information) and disadvantages (loss of information) should be studied together. Although the present work involves arrays of multifeatured objects to be recognized, for the sake of easy visualization we first present the overcompression idea in the context of sequential stimuli. For instance, if participants make use of a compression process to recode information, capacity should be increased in situations where patterns exist (for instance, for retaining a display set such as  $\square \square \square$  sequentially as shown, as the compact description " $2\blacksquare\Box$ "). This description would be sufficient to recognize the same sequence reproduced as a probe display or reject a sequence in which something has changed from the studied sequence. However, the putative compression process could also lead to specific errors when patterns are badly (or overly) compressed. If a participant studied the display set  $\blacksquare \Box \blacklozenge \Box$ it might be recoded as the rule: " $2\blacksquare\Box$ , but rotate second dark square." However, the qualification about rotating the second dark square could be lost from working memory, so, if the test display were  $\blacksquare \Box \blacksquare \Box$ , the participant would incorrectly find a perfect match to the now overcompressed or *lossy* representation  $2\blacksquare\Box$ .

We presume that a poorly managed compression process could distort perceptual organization instead of producing benefits. The above-mentioned studies have mostly insisted on the benefits of compressibility. Our hypothesis in the present study is that compression in working memory could also be detrimental to the recall process. The reason is that we can expect the memory content to be over-simplified if memorization is driven by a compression process that seeks to reduce information load. Also, simplification errors should depend on compressibility levels, as less random errors should be expected when stimulus sets are more structured. We should thus expect a greater number of compression errors with greater compressibility.

Our stimulus displays were made of objects comprising a variable number of conjunctions of features, to manipulate the complexity of the display. Complexity was manipulated to allow recoding of information based on the compressibility metric defined by Feldman (2003), which is adapted to Boolean dimensions (i.e., dimensions made of two discrete features). Feldman (2003) described complexity by using spatial dimensions represented by hypercubes, also as Hasse diagrams. A hypercube is just an extension of a square representing four two-dimensional objects or a cube representing eight three-dimensional objects. These diagrams are useful to represent the similarities between objects. In a cube, two objects related by an edge share only one feature. The number of edges that one needs to follow to join two objects represent the number of dissimilarities between two objects. A fourdimensional hypercube represents two joined threedimensional cubes, so one needs to follow three edges and then switch cubes to find the object differing in four features from one object. By representing the locations of some objects using black dots in a hypercube allows one to represent the relational structure between objects. Figure 1 shows the similarity structure of the stimulus items (both with and without the probe) for each of our experimental conditions. To represent how this similarity structure can be best represented using a minimal number of features, the compressibility metric of Feldman is based on the minimal disjunctive normal forms corresponding to the selected features for a given structure in a hypercube (Table 1). In Fig. 1 the complexity number indicates the minimal number of features that describe the chosen subset of stimuli, the stimuli are marked with a dot in each of the hypercubes, and the structure column always contains a new dot in the hypercube, which represents the chosen probe. For instance, the hypercube in the first row and third column of Fig. 1 can be summarized by the spatial rule "any object that is on the *left* and at the bottom." When using the reference set of features at the

top of Fig. 1, this spatial rule corresponds to the feature rule "any object that has a *left-disc* and that is *square*." Because the rule mentions two literals (i.e., the two features just mentioned), the complexity of the rule sums to 2. Another example is the hypercube in the first row and first column of Fig. 1 can be summarized by the spatial rule "any object on the left and at the bottom (except the object in the *back* of the *right-cube*)." Because the rule mentions four literals (i.e., the four traits just written in italics), the complexity of the rule sums to 4. Using the specific dimensions we used in our experiment, the rule would be exemplified as: "any object that is on the leftdisc and square (except the hatched and red one)" (but note that the experiment actually used rotations of the dimensions, so the features for a given trial would be randomized, although obeying the specific structure). One more complex example would be the last row of the first column, summarized by "any object on the left and at the bottom (either on the front of the left-cube or at the *back* of the *right-cube*), plus the object on the *top* right corner on the *front* of the *right-cube*." The sum of the traits in italics is here 10. One implementation would be: "square with left-disc (either the plain and blue or the red and hatched), plus the plain red circle with a rightdisc."

In the present study we aimed at predicting accuracy and response times (RTs) in a comprehensive way as a function of stimulus-set complexity. Our experiments allowed us to study memory for diverse study sets of three objects presented simultaneously, followed by a test object that participants were asked to categorize as new (absent from the array) or old (present in the array), as exemplified in Fig. 2. Our general method consisted of displaying arrays of three objects that were designed to allow associations to be formed. The objects were four-dimensional stimuli varying in shape, color, texture (plain vs. hatched), and direction of a white dot within the shape (left vs. right).

This method allowed us to study how introduction of a lure could modify the participants' representation. Consider, for example, the array to be studied in Fig. 2. Given that two of three discs appear on the left of the objects (i.e., for the two squares), an over-compressed representation might include the disc on the left in all three object representations. In that case, a probe item that was a blue circle with the disc on the left would be incorrectly judged to be present in the array. As another example based on Fig. 2, the probe item shown -a red circle with the disc on the right - could be incorrectly identified based on an over-regularized representation in which blue was assigned to all squares, and red to all circles. However, the latter representation is less plausible than the former because both "red" and "dot on the right" are statistically under-represented in the study display. By studying

Table 1	Minimal formulae for conditions used in Experiment 1 and
Experimen	nt 2 based on the catalog of Feldman (2003). For instance, the
features a-	a', b-b', c-c', and d-d' represent the features Blue-Red, Square-

Circle, Plain-Hatched, and Left-disc-Right-disc, respectively, as shown in Fig. 1. Correspondence between letters and features was randomly assigned in each trial

IFS Code	Initial complexity	Minimal formula	Final complexity	Minimal formula	Change in complexity	Sum of shared features	Exp.
4-2-8	4	a'b'(cd)'	2	a'b'	-2	8	2
4-7-7			7	a'(b'(cd)'+bc'd')	+3	7	2
4-8-2			8	a'b'(cd)'+abcd	+4	2	1
4-8-4			8	a'b'(cd)'+abc'd'	+4	4	1 and 2
6-5-9	6	a'(b'c'+bcd')	5	a'(bc+b'c')	-1	9	2
6-6-4			6	bcd'+a'b'c'	0	4	1
6-9-3			9	bc(a'd'+ad)+a'b'c'	+3	3	1 and 2
6-9-5			9	cd'(a'b+ab')+a'b'c'	+3	5	1
6-10-4			10	a'(b'c'+bcd')+ab'cd	+4	4	1 and 2
7-6-4	7	a'b'c'+abcd'	6	abc+a'b'c'	-1	4	1
10-8-4	10	a'b'(c'd' + cd) + abc'd'	8	(c'd'+cd)(ab+a'b')	-2	4	1 and 2
10-12-4	10	a'b'(c'd'+cd)abc'd	12	ab(c'd+cd')+a'b'(c'd'+ cd)	+2	4	1 and 2

Note. IFS = Initial-Final-Sum of shared features. All conditions are displayed visually in the Online Supplementary Material

various conditions similar to this example, we discuss the advantage of the present method, which we believe goes beyond previous research that mostly focused on the benefits of compression processes. Also, the present study goes a step further by investigating how both errors and RTs could result from the putative compression process.

# **Experiment 1**

#### **Overview**

The aim of Experiment 1 was to study immediate memory for visual scenes made of arrays of three multi-dimensional objects presented simultaneously. Each array was followed by a probe and the task consisted of deciding as quickly as possible whether the probe was new or old. The design for a given trial is exemplified in Fig. 2. Based on a compressibility metric we develop in this section, our goal was to assess the quality of the participants' representations using probes that, depending on the conditions, could completely or partially match the three stimulus objects of the memory array. We recorded RTs and accuracy to measure the effect of the compressibility of the memory sets. Interestingly, our method also consisted of characterizing the effect of the probe interacting with the memory set, by describing whether introduction of the probe could fool the participants. The to-be-tested idea was that the

participants' representations could be prone to be modified upon introduction of the probe.

# Method

#### Participants

Thirty French participants ( $M_{age} = 31.8$  years; SD = 10.6) volunteered to take part in the experiment. The sample included 24 females and six males, having completed between 0 and 8 years of higher education. The experiment was approved by the local ethics committee (CERNI) of the Université Côte d'Azur and the experiment was conducted with the informed consent of the participants. To estimate our minimal sample size, we referred to the study by Feldman (2000) in which 45 subjects were asked to memorize similar sets of three fourdimensional Boolean objects. Feldman achieved sufficient power to show a relationship between complexity and proportion of correct recall ( $R^2 = .37$ ; a similar positive relationship was shown with three three-dimensional objects using 22 participants,  $R^2 = .98$ ). In the study by Feldman though, the participants had to observe for 20 s the 16 four-dimensional objects (the four to-be-memorized objects called "positive examples" appeared in the upper half and all other objects appeared in the lower half). The participants were then asked to categorize all 16 objects presented randomly as positive or negative during a block called the "categorization task." This

	Reference set of features for Exp. 1 and Exp. 2								
Initial structure	Initial complexity	Final structure (initial set + probe)	Final complexity	Change in Complexity	Sum of shared features between the initial set and the probe				
ØØ	4		2	-2	8				
	4	II	7	+3	7				
	4	Į	8	+4	2				
<b>I</b> I	4		8	+4	4				
Í	6	μŢ	5	-1	9				
ſſſ	6	ÍÍ	6	0	4				
Í	6	μſ	9	+3	3				
Í	6		9	+3	5				
Í	6	ШП.	10	+4	4				
Í	7	ÍÍ	6	-1	4				
Í	10		8	-2	4				
	10	<u>I</u> I	12	+2	4				

Fig. 1 Initial structures and complexity estimates, with and without the probe. Change in Complexity corresponds to the evolution of the complexity between Initial structure and Final structure. The sum of

shared features represents the number of features that a probe (lure) shares with the stimuli of the initial set

categorization task seemed more demanding than our memorization task, so we estimated roughly that a sample of 30 subjects would at least allow us to observe the Initial complexity effect.

#### Dimensionality and compressibility of stimulus sets

The task was created using the Python library PsychoPy (Peirce, 2007). The four-dimensional stimuli of Kibbe and Kowler (2011) were used to draw the individual stimulus items based on different shapes (circle vs. square), color (blue vs. red), texture (plain vs. hatched), and disc position (left vs. right).

The combinations of features allowed the construction of 16 possible objects. We used these stimuli because our experiment required four-dimensional stimuli not varying in size in order to have equal distances between stimuli. Also, we chose to use displays of three objects, because this cardinality facilitated equalization of distances within arrays; all objects were equidistant in the array, as they were arranged as points of an equilateral triangle. Each trial used three different objects shown simultaneously on a single display on a white background (Fig. 2). Table 1 shows all of the combinations that were used in the task to generate the trials, based on the complexity of the display structure and the complexity of the structure including the probe. Following introduction of the probe, our manipulation made the complexity of the four items altogether (i.e., the three stimuli plus the probe) increase, decrease, or stay constant.

The second column of Fig. 1 indicates the complexity of the initial set displayed in the first column (in the first column, the chosen stimuli are marked with a dot in each of the hypercubes). For instance, the last stimulus set has a complexity of 10 because ten features are necessary to minimally describe



Fig. 2 Timeline of a single trial in the main experiment

the entire set. The third column indicates the complexity of the final stimulus set (initial set to which we added the probe/ lure). For instance, the last stimulus case of the table has a complexity of 12 because the entire set of four stimuli is not easily compressible, and in that case the adjunction of the probe makes the entire structure more complex than the initial structure of the three stimuli. The last column of Fig. 1 also indicates the sum of shared features between the initial structure and a lure. For instance, the sum of shared features is easily visible in Fig. 3a where there are only two shared features between the initial structure on the left and the probe on the right (the feature *red*, and the feature *hatched*). In Fig. 3b, there are, however, four shared features between the initial structure on the left and the probe on the right. The number of shared features served as a way to double the number of observations for each case of interest, by taking advantage of all of the possible variations that were allowed.

The reason why we restricted the choice of structures and probes was to manipulate two independent variables independently. We used a total of nine possible structures listed in Table 1. Each case was numbered using the values described in Fig. 1 for each variable, using a triplet for the three respective measures (initial complexity, final complexity, and sum of shared features). For instance, in Fig. 3a and Fig. 3b, the structures are, respectively, 4-8-2 and 10-8-4. The structures 4-8-2, 4-8-4, 6-6-4, 6-9-3, 6-9-5, 6-10-4, 7-6-4, 10-8-4, 10-12-4 are shown in Fig. 1 and recapitulated in Table 1. The number of features between the initial set and the probe was considered a control variable.

The combinations of different initial structures and different probes allowed us to generate two main independent variables: (1) Initial/display set complexity (inversely corresponding to the compressibility of the display set), (2) Change between Initial Set Complexity and Final set complexity (inversely corresponding to the compressibility of the stimuli of the display set and the probe combined). For instance, in Fig. 3a, the initial structure presents a complexity of 4, while the final structure presenting a complexity of 8 led to an increase of four points in complexity (Change = 4). Also, note because that there are two features in common between the initial structure and the probe in the case of Fig. 3a, we code this case using the triplet 4-8-2 (corresponding to, respectively, Initial Set Complexity, Final set complexity, and Number of shared features between the initial set and the probe). Another example is 10-8-4, in which the change in complexity is equal to 8 - 10 = -2, which reflects the idea that the probe can decrease the complexity of the initial structure.

Two predictions follow from the two main factors Initial Complexity and Change in Complexity:

 A low complexity of the initial set was hypothesized to account for greater memory performance in general (both



**Fig. 3** Sample of initial sets (stimulus triplet on the left side) and probes (singleton stimulus on the right side). (a) The initial structure presents a complexity of 4, while the final structure presents a complexity of 8. There are two features in common between the initial structure and the

change and no-change trials can be performed more easily when the initial set can be potentially compressed, with faster response times considering that a more compressed representation could be rapidly decompressed); conversely, performance was expected to be lower when the complexity of the initial set is higher (as participants cannot find a way to simplify information, information load is higher, and decompression time takes longer).

2) Introduction of the probe (i.e., a fourth element) should interact with the initial set of three objects. In particular, the probe was hypothesized to deceive the participant when the complexity (of the four elements taken together) decreases in comparison to that of the initial set (i.e., a negative Change) because the probe fitted well to the display set; conversely, it was hypothesized that the participant would better detect a change trial when the complexity increases because an increase of complexity assumed a very different probe from the display set. This second less intuitive prediction is still based on the compression account: participants can be deceived by a lure if they over-compress information. One simple example would be a presentation of three colored squares (white, dark grey, black): if not correctly encoded (for instance based on a lossy compression such as "all squares" or "squares from white to dark"), the participant would have a greater chance to falsely recognize a light grey square,

probe. This case is coded 4-8-2. (b) Second case: The initial structure presents a complexity of 10, while the final structure presents a complexity of 8. There are four features in common between the initial structure and the probe. We therefore coded this case 10-8-4

but probably not a circle that would involve a larger modification of complexity. We thus expected that a high increase in complexity due to introduction of the probe would have a greater chance of not fitting a compressed or over-compressed representation. The factor Number-ofshared-features was not of direct interest in the present study and it was only thought to better generalize our results. To elaborate on an example based on our material, let us presume a participant can have trouble encoding the display set (three first items only) in Fig. 3a as "anything that is not-right-disc and not-circle (but not hatched and red in the same time)." Note that it should be assumed that the participants are not expected to encode the set more minimally such as using the description "not-right-disc and not-circle" (or "square and left-disc") because they are instructed to perform successive trials in which all eight potential features can be a determinant for a correct response. Taking into account this constraint, when trying to optimize the storage process using the available compressibility, the participant could encode the stimulus set as "square and left-disc" and leave out the exception. In that case, the introduction of the probe square with red hatching would satisfy the over-compressed representation. A false alarm could therefore be committed when the participant is required to decide whether the probe was present or not in the original stimulus set. When the probe corresponds to one of the three displayed objects, the probability of an omission would be very low based on the over-compressed participant's representation. However, if the probe was a circle, the three squares of Fig. 3a mixed with the circle would not produce a new cohesive ensemble. The complexity of the new ensemble would not fit the one of the initial set of three stimuli, so rejection of the probe would be facilitated.

#### Procedure

Participants sat approximately 60 cm from the computer screen. Participants were required to enter their response for each trial using two keys of a numeric keypad (keys 4 and 6) depending on whether they judged the probe had been previously presented in the stimulus display or not. After receiving a series of basic instructions to complete the task, participants started the experiment with a series of 15 trials as a warmup, including feedback. Next, the participants were administered 540 trials with no feedback. Within each of the nine experimental structures, the probe in the no-change trials corresponded to one of the three previously presented objects for 30 trials. There were 30 other change trials with the probe (i.e., a lure) absent from the stimulus display. Participants could take a quick break after the first 200 trials, and after 400 trials. The 540 trials were permuted for each participant.

Each trial began with a 2,000-ms fixation cross followed by a 500-ms blank screen (white). The stimulus displays then appeared for 1,000 ms, followed by a second 500-ms blank screen (white). Note that the stimulus display was not followed by a random mask to allow maximal encoding of the stimuli. Then, the probe was shown for 300 ms before a final blank screen (white) that allowed sufficient time for participants to enter their responses (see Fig. 2). The next trial was initiated automatically by the program. We measured the number of hits (i.e., the participant recognized the probe as one of the stimulus items of the stimulus display), false alarms (FAs), omissions, and correct rejects (CRs) and RTs.

#### Results

The present analysis focuses on two aspects of participants' performance. For each of the following predictions, we ran a linear mixed-effects model using participants as a random variable and allowing different intercepts and slopes across participants (see Brown, 2020; Singmann & Kellen, 2019). Each model used only one dependent variable per analysis, either drift rates or FA rates. Drift rates allowed us to combine the proportion of correct responses, mean correct RTs, and their variance. Drift rates is a parameter of the diffusion model (Ratcliff, 1978) that has been developed for speeded binary decision processes. This parameter generally described as one

of the most relevant of the model allows one to describe how fast the decision reaches one of two optional responses. Drift rates allowed obtaining an overview of correct performance, instead of using the three dependent variable hit rates, hit RTs, and correct rejection RTs (see Online Supplementary Material (OSM), which provides the details for these variables). Because our data set is rather small and we used a twoalternative forced-choice task (and to avoid the complex parameter-fitting procedure of the original Ratcliff diffusion model), drift rates were estimated at a macroscopic level using the EZ diffusion model (Wagenmakers et al., 2007) and were calculated using the R package EZ2. Beforehand, we checked that RTs were right-skewed, that skewness was more pronounced with increasing complexity, that RTs were comparable for each type of response (correct and incorrect), and that RTs were comparable for all incorrect responses through subjects and conditions.

Initial Complexity and Change of Complexity were considered the main independent, manipulated factors. We also used the number of shared features between the probes and the display set as a covariate to better control for similarity effects. We also computed interactions between factors to refine our analyses but we made sure our models were not overly complex by providing values of the Akaike Information Criterion (AIC; see Akaike, 1987), which offers a trade-off between the goodness of fit of the model and its simplicity.

Firstly, we posited that higher complexity of the initial set should account for lower memory performance. We thus expected lower drift rates and higher FA rates because of a failure to encode stimulus features of the highest complexity sets.

Secondly, an increase in complexity caused by the introduction of a lure probe should account for greater memory performance (lower FA rates and higher drift rates). The rationale was that an increase in complexity due to introduction of the lure probe has a greater chance of not fitting a compressed (or over-compressed) representation of the display set. The decision (i.e., reject the lure) was thus expected to be easier with a greater change in complexity. Conversely, we expected higher FA rates and lower drift rates when introduction of the lure does not increase complexity, with the idea that this manipulation can help detect an errant compression process.

For more detailed descriptive statistics, the OSM shows the descriptive statistics of the hit rates, FA rates, hit RTs, and correct rejection RTs across the paired structures.

We only removed 2.5% of the data corresponding to RTs less than 250 ms and greater than 2,000 ms. We chose to keep relatively long RTs (skewness = 1.23) to study the potential effect of complexity.

For each of the two dependent variables, the following subsections first describe the result of the mixed model applied to a given dependent variable before presenting the statistical tests for each of the paired comparisons. For the paired comparisons, we applied the Bonferroni correction for multiple paired comparisons (each correction was applied within a given dependent variable, therefore never exceeding two comparisons, thus with a threshold at p = .025). Tables 2 and 3 show the results of the linear mixed models run on drift rates and FA rates. The paired comparisons across all dependent variables are shown in Fig. 4.

#### Drift rates

To analyze whether complexity levels affected memorization, we first ran a linear mixed model to study the influence of Initial Complexity and Change of complexity on drift rates. We also use the method of paired comparisons when the factor Sum of shared features could be maintained constant.

Table 2 shows the results of the linear mixed model run on drift rates as a function of Initial Set Complexity, Change in Complexity, and adding Sum of shared features as a control factor. The Change in Complexity factor could have positive values (i.e., complexity increased when introducing the lure), negative values (i.e., complexity decreased when introducing the lure), or a null value (complexity did not change in spite of the new structure). See Table 1.

Contrary to our expectations, the mixed model showed no significant decrease of drift rates as a function of Initial Set Complexity. We nevertheless observed an increase of drift rates with a higher change in complexity between the initial set and the final set (t(2.258) = 2.905, p = .004), as expected, but this effect was counterbalanced when complexity of the initial set was the highest (which is captured by the interaction between the two factors:  $t(2.258) = -2.220, p \le .027$ ). Nevertheless, a decrease was observed for the two additional paired comparisons allowing testing of the effect of Initial set complexity (Fig. 4a): the drift rates decreased significantly for both the pairs 4-8-4 versus 10-8-4 (t(29.00) = -6.48, p < .001) and **6**-6-4 versus 7-6-4 (t(28.95) = -2.56, p = .016). On the other hand, the analysis based on the comparisons 6-6-4 versus 6-10-4 (in the condition 6-6-4 the probe did not change the complexity, contrary to in the condition 6-10-4 in which the probe increased the complexity by 4 points) and 10-8-4 versus

 Table 3
 General effect of Initial set complexity, Change in complexity, and Sum of shared features on false alarm rates

	Estimate	Std. error	df	t value	p value
Intercept	2.131	6.99	6.603	3.046	.002 **
Initial Set Complexity	-2.073	1.37	7.505	-1.509	.131
Change in Complexity	-3.542	4.97	1.862	-7.127	.000 ***
Sum of shared features	-2.372	1.70	7.349	-1.400	.162
Initial Set Complexity × Change in Complexity	3.729	6.36	7.731	5.865	.000 ***
Initial Set Complexity × Sum of shared features	3.865	3.40	7.727	1.139	.255

**10-12-4** (respectively a Change of -2 and +2) were conducted. In Fig. 4c, paired comparisons did not show a significant effect (6-6-4 vs. 6-10-4, t(2.900) = .59, p = .059; 10-8-4 vs. 10-12-4, t(2.900) = -0.58, p = .565). Drift rates corresponded to a good answer rate and good answer RTs but these two last paired comparisons were interested in the effect of the modification of the complexity by the lure probe; however, the drift rate seems to be an insufficient measure to capture this effect. Therefore, it was chosen to observe FA rates independently.

#### False alarm (FA) rates

The following two analyses concern situations in which a lure was displayed, allowing us to study both Initial Complexity and Change of Complexity. For FA rates, we ran a linear mixed model including an interaction term between the two factors of interest, and we also ran paired comparisons as a function of the factor Change of Complexity (by maintaining the two other factors constant).

Table 3 shows the result of the linear mixed model run on FA rates as a function of Initial Set Complexity, Change in Complexity, and adding Sum of shared features as a control factor. A FA corresponds to an incorrect answer in a change trial (a false recognition of the lure as part of the initial display set).

Contrary to our expectations, the mixed model showed no significant decrease of FA rates as a function of Initial set. We

Table 2         General effect of Initial set complexity, Change in complexity, and Sum of shared features on dr	ift rates
--	-----------

	Estimate	Std. error	df	t value	p value
Intercept	8.807	2.11	2.295	4.168	.000 ***
Initial Set Complexity	-6.928	4.35	2.269	-1.592	.112
Change in Complexity	4.192	1.44	2.258	2.905	.004 **
Sum of shared features	-4.759	5.43	2.258	-0.875	.382
Initial Set Complexity × Change in Complexity	-4.319	1.94	2.258	-2.220	.027 *
Initial Set Complexity × Sum of shared features	1.310	1.12	2.258	1.169	.243

# Effect of Initial set complexity



Effect of Change in complexity



Fig. 4 Effects of Initial complexity and Change of complexity in Experiment 1. By panels, effects of the Initial set complexity as measured by (a) drift rates (×100), and (b) false alarm rates, and effect of Change in complexity as measured by (c) drift rates (×100), and (d) false alarm rates. The plots are restricted to the paired comparisons ceteris paribus. Error bars represent ±1 SE. *Note.* The triplets of numbers used to code conditions correspond to, respectively, Initial Set Complexity, Final set complexity, and Number of shared features between the initial set and the probe

nevertheless observed a decrease of FA rates with a higher change in complexity between the initial set and the final set (t(1.862) = -7.127, p < .001), as expected, but this effect was counterbalanced when complexity of the initial set was the highest (which is captured by the interaction between the two factors: t(7.731) = 5.865, p < .001). This result is apparent in the analysis based on the comparisons 6-6-4 versus 6-10-4 and 10-8-4 versus 10-12-4; in Fig. 4b, paired comparisons showed a significant increase in FAs with an increased Initial Set Complexity (4-8-4 vs. 10-8-4, t(1.733) = 3.579, p < .01; 6-6-4 vs. 7-6-4, t(1721) = 3.135, p = .002) and in Fig. 4d, paired comparisons showed asignificant decrease in FAs with an increased change in complexity when Initial Set Complexity was low (**6-6**-4 vs. **6-10**-4, t(1.742) = -3.494, p < .001) but not when Initial Set Complexity was higher (10-8-4 vs. **10-12**-4, t(1.712) = 1.72, p = .086). For the latter comparison, the greater complexity of the initial set (i.e., 10) might not have facilitated memorization whatsoever (we observed the lowest hit rates for these two structures as shown in the preceding analysis).

To go further with our analysis on FA rates, we attempted to control the effect of the number of new features caused by the introduction of the probe (features in the lure not presented in the initial set). Indeed, if participants only saw a set of squares in the initial set before being presented with a circle in the display test, this new feature "circle" could allow an easier rejection of the lure (Mewhort & Johns, 2000), which could account for a decrease in the FA rate. The emergence of one or two new features in the lure effectively decreased significantly the FA rates in comparison to the conditions in which no new feature appeared (F(1,29.058) = 25.31, p <.001). Nevertheless, when we only selected the conditions for which no new feature appeared in the lure, we still found a significant effect of Initial Set Complexity (F(1,445) =20.13, p < .001). There was, however, no effect of Change in complexity (F(1,32) = 1.59, p = .216), but note that the conditions having no new features in the probe were the three conditions with the highest Initial set Complexity and previous analyses indicated that no effect of Change in complexity appeared when the initial complexity was too high. Thus, effects observed on FA rates were not entirely linked to the number of new features introduced with the lure.

# Comparative analysis of drift rates and FA rates for all independent variables

Our last analysis was based on the two dependent variables (drift rates and FA rates), which were both predicted to be influenced by the two main factors (Initial Set Complexity and Change in Complexity). Our goal was to verify whether the full model including the interaction term was effectively the most parsimonious model. The two mixed-model analyses progressively included the two factors of interest and their interaction, and the additional control factor Shared features. The successive models were tested to obtain the most parsimonious account of the data, by computing the AIC for each model (see Table 4). A lower AIC corresponds to a more optimal model (i.e., considering that a parsimonious model leads to an adequate fit associated with minimal model complexity). In the case of drift rates, the mixed model including the three factors led to the best AIC (i.e., -2618.618). For the FA rates, the best model corresponded to the model including only Sum of shared features (i.e., -1795.354).

#### Discussion

Our aim was to test whether visual memory capacity is determined by a process of data compression by studying correct responses, errors, and RTs. Our compressibility metric was based on the complexity of the stimulus material, with more complex material being theoretically less compressible. We measured accuracy and speed of responses in order to analyse FA rates and drift rates. The drift rates integrated accuracy and RTs for correct responses. Our assumption was that individuals can develop on the spot a compressed representation of a display set obeying regularities (e.g., correlated features), and this design was thought to test the compression hypothesis without the need to retrieve pre-existing chunks (as is the case in the studies by Brady et al., 2009, and Reder et al., 2016). We also assumed that the compression process is a trade-off, with the risk of compressing too much information resulting in lossy representations. We expected that the presence of compressibility in the display set would enable greater memory performance and faster performance based on the idea that shorter compressed representation would take less time to be decompressed. A more novel prediction was that the introduction of a lure would interact with the compressed representation to modify the initial perceived complexity.

For the factor Initial Set Complexity (i.e., the complexity of the display set), as predicted, our findings showed a decrease in performance (smaller drift rates) with higher complexity, indicating a higher cognitive load for less compressible display sets in pairwise comparisons of the complexity structures. The effect of complexity on FA rates was not found in the mixed model, but it was detected in one paired comparison.

Model	Variables	AIC	
		Drift rates	FA rates
M1	Initial Set Complexity	-2712.978	-1823.150
M2	Change in Complexity	-2692.762	-1844.364
M3	Sum of shared features	-2670.019	-1795.354
M4	Initial Set Complexity Change in Complexity	-2678.066	-1866.230
M5	Initial Set Complexity Sum of shared features	-2675.376	-1802.032
M6	Initial Set Complexity Change in Complexity Sum of shared features	-2618.618	-1844.089

Table 4 Comparisons of mixed model based on the AIC criterion for Drift rates and False alarm (FA) rates

The numbers in bold and underlined correspond to the most optimal model for each dependent variable (lowest AIC), i.e., M6 (including the three variables studied) for Drift rates and M3 (including only the variable Sum of Shared Features) for FA rates

When we manipulated a change in complexity with introduction of a lure, our findings showed less pronounced effects. Within the paired comparisons, we observed a large decrease of FAs with a higher change in complexity when the complexity of the initial set was initially low (i.e., in the pair 6-6-4 vs. 6-10-4). We did not observe the expected effect when the complexity of the initial set was initially high (i.e., in the pair 10-8-4 vs. 10-12-4). Our interpretation is that the lure might not affect the decision process when encoding of the stimulus set was already too difficult. Analyses on drift rates, however, showed the expected effect of Change in Complexity in the mixed model: a higher change in complexity resulted in higher drift rates, meaning that a probe can easily be identified as not belonging to the initial set when complexity changes. Change in Complexity nevertheless interacted with the Initial Set Complexity factor, again revealing that a Change in Complexity did not occur when the complexity of the initial set was initially high (10-8-4 vs. 10-12-4). Aside from the general effect captured by the mixed model, no significant direct effect could be isolated from the paired comparisons.

Finally, the effect of new features in the probe was tested and an effect of Initial Set Complexity on FA rates was still found even when no new feature was introduced in the probe.

# **Experiment 2**

Experiment 1 suggested an effect of change in complexity caused by the addition of a probe to the initial stimulus set. This effect could be due to a modified perception of the initial complexity of the stimulus set when the probe has to be compared to the memorized stimuli. However, this effect was only observed in the paired condition **6-6**-4 versus **6-10**-4, with lower performance in **6-6**-4 as participants seemed to be more lured by a probe that did not modify the level of complexity;

this difference was not observed in the more complex pair of conditions **10-8**-4 versus **10-12**-4. We thus concluded in Experiment 1 that the effect of change in complexity primarily depends on initial stimulus set complexity, as participants might not be able to properly encode information when complexity was too high, but the experimental conditions did not allow to study intermediate levels of complexity.

In Experiment 2, new conditions were tested to study a larger and better controlled range of complexity effects. The new conditions were thought to better observe the two opposite effects of change in complexity (i.e., Increase in complexity with introduction of the probe vs. Decrease in complexity with introduction of the probe) for different levels of initial set complexity (Low vs. Medium vs. High, corresponding to initial stimulus set complexity equal to 4, 6, and 10, respectively).

#### Method

#### Participants

Forty French participants ( $M_{age} = 25.9$  years; SD = 6) volunteered to take part in this experiment. The sample included 25 females and 15 males, having completed between 0 and 5 years of higher education, who were recruited at Université Côte d'Azur or from the Alpes-Maritimes county through advertisements. We targeted a sample roughly equivalent to that of Experiment 1, but our recruitment efforts yielded some additional participants who registered.

#### Procedure and materials

The procedure of Experiment 2 was similar to Experiment1, except that we focused on a more comprehensive ensemble of complexity levels. Three levels of initial set complexity were selected: low (4), medium (6), and high (10). For each level, the probe could either decrease or increase the complexity of the initial set (see Table 1).

#### Results

As in Experiment 1, the analysis first focused on the effect of Initial Set Complexity on drift rates, which were hypothesized to reflect better memory performance with lower initial complexity. As in Experiment 1, we checked that RTs were rightskewed and the skew became more pronounced with the increase of complexity, that RTs were comparable for each type of response (correct and incorrect), and that RTs were comparable for all incorrect responses through subjects and conditions. The second hypothesis again posited that a decrease in complexity caused by the introduction of the lure probe should account for lower memory performance (higher FA rates). Again, the rationale of this last prediction is that if participants over-compress information of the initial stimulus set, there is a greater chance that the participant can be lured by a probe that makes the complexity of the four objects decrease.

To use a range of RTs similar to Experiment 1, we removed 7% of the data corresponding to RTs less than 250 ms and greater than 2,000 ms. The Appendix shows the descriptive statistics of the hit rates, FA rates, hit RTs (i.e., the time taken to recognize the probe as being part of the stimulus items), and correct rejection RTs across the conditions (classed by level of initial complexity). As in Experiment 1, all linear mixed-effects models were run using participants as a random effect with individual intercepts and slopes.

#### Drift rates

Figure 5a shows the result of the drift rates as a function of the Initial Set Complexity. As in Experiment 1, the linear mixed models testing the unique effect of the factor Initial Set Complexity showed a significant decrease of drift rates (F(1,235) = 12.031, p < .001) with a higher complexity.

Figure 5b shows the drift rates as a function of Complexity of the initial set and Change of Complexity. Again, as in Experimenet 1, the expected increase of performance with the highest initial complexity was not observed when complexity increased with the introduction of the probe. A striking result was the effect of the type of probes on drift rates when complexity of the initial set was low. In this case, performance was the weakest when introduction of a lure decreased the complexity of an already simple initial display set, whereas performance was the highest when presentation of the lure increased complexity. Paired comparisons also showed a decrease of drift rates with an increase of Initial Set Complexity for both paired conditions (4-8-4 vs. 10-8-4, t(3.900) = -6.04, p < .001; 10-8-4 vs. 10-12-4, t(3.900) = -2.12, p = .041). Table 5 shows the results of the linear mixed model run on drift rates for the factors Initial Set Complexity and Change in Complexity (adding Sum of shared features as a control factor), confirming all main effects and interactions.

Figure 5d shows the combined effect of both Initial Set Complexity and Change in Complexity on memory performance. The combined effect was independent of the factor Shared features, which suggests that complexity effects were not based solely on similarity effects between the probe and the stimulus set.

#### FA rates

Like in Experiment 1, the following analysis on FA rates concern situations in which a lure was displayed, allowing us to study the combined effect of Initial Complexity and Change of Complexity and their interaction.

Figure 5c shows the effect of Change in Complexity according on the factor Initial Set Complexity on FA rates. Consistent with our assumptions, the participants correctly rejected the probe when it increased the complexity of the initial stimulus set, with no effect of the initial stimulus set complexity. On the contrary, the participants faced difficulty in rejecting the probe as a function of the level of Initial Set Complexity when complexity decreased with the probe. Table 6 shows the result of the linear mixed model run on FA rates as a function of Initial Set Complexity and Change in Complexity (again, adding Sum of shared features as a control factor). The results are similar to those of Experiment 1, as the mixed model showed an interaction indicating a different effect of Change in Complexity as a function of Initial Set Complexity. The analysis also indicated a significant interaction between Initial Set Complexity and Shared features. Nevertheless, Fig. 5e shows a combined effect of both Initial Set Complexity and Change in Complexity on performance independent of the factor Shared features, which suggests that complexity effects are not based solely on similarity effects between the probe and the stimulus set. Also, it is interesting to see that the case 10-8-4 for which Initial Complexity was the highest (thus making the initial set difficult to memorize) and the factor Change in Complexity the lowest (i.e., minus two, thus making the lure deceitful in case of an over-compression process) produced the highest FA rate. But overall, the clearest decreasing trend was observed from 6-10-4 to 4-8-4, and because the Change in Complexity is similar in these two cases, initial complexity seems to take precedence in accounting for the decreasing trend. Moreover, paired comparisons also indicate an effect of Initial Set Complexity on FA rates with the increase of FA rates with the increase of Initial Complexity (4-8-4 vs. 10-8-4, t(2.230) = 5.57, p < .001) but no effect of change in complexity when the Initial Set Complexity was too high (10-8-4 vs. 10-12-4, t(2.199) = -1.11, p = .269).



◄ Fig. 5 Effects of Initial complexity and Change in complexity in Experiment 2. By panel, (a) Initial Set Complexity measured by the drift rates (×100), (b) Initial Set Complexity and change in complexity measured by drift rates (×100), (c) Initial Set Complexity and change in complexity measured by false alarm (FA) rates, (d) Initial Set Complexity measured by drift rates (×100) with fixed Shared Features, and (e) Initial Set Complexity measured by FA rates with fixed Shared Features. Error bars represent ±1 SE. *Note.* For figures d and e, the triplets of number used to code conditions correspond to, respectively, Initial Set Complexity, Final set complexity, and Number of shared features between the initial set and the probe

To refine these results on FA rates, we controlled the effect of the number of new features with introduction of the probe (i.e., the features in the lure that were not part of the initial set) as in Experiment 1. The apparition of one or two new features in the lure effectively decreased significantly the FA rates compared to the conditions in which no new feature appeared (F(1,38.597) = 28.60, p < .001). Nevertheless, when we only selected the conditions for which no new feature appeared in the lure, we found both a significant effect of Change in Complexity (F(1,119) = 21.17, p < .001) and an interaction between Initial Set Complexity and Change in Complexity (F(1,112) = 20.14, p < .001). Thus, effects observed on FA rates were not entirely linked to the number of new features introduced with the lure.

Also, the case of the condition 4-2-8 deserves attention, as it is the only condition presenting a low initial complexity and a decreased change in complexity, but more importantly, the FA rate for this condition was the highest observed in our experiments. To refine the analysis of this condition, we split the data according to block number in the task. We only included the conditions with zero new features introduced with the probe, as well as conditions in which complexity decreased. The data were split as a function of the first block (first decile of trials) versus the rest of the blocks collapsed, for simplicity purposes. Figure 6 shows that FA rates were the highest in the first block for the condition 4-2-8 in comparison to the rest of

**Table 5**General effect of Initial Set Complexity, Change inComplexity, and Sum of shared features on drift rates shown in Fig. 5b

	Estimate	Std. error	df	t value	p value
Intercept Initial Set Complexity Change in Complexity Sum of shared features Initial Set Complexity ×	8.97 -5.23 4.88 -6.45 -6.03	1.05 1.50 9.76 1.82 1.26	2112 2109 2313 2120 2064	8.112 -3.484 4.997 -3.544 -4.801	.000 *** .000 *** .000 *** .000 ***
Change in Complexity Initial Set Complexity × Sum of shared features	9.54	2.92	2064	3.261	.001 **

**Table 6** General effect of Initial Set Complexity, Change inComplexity, and Sum of shared features on false alarm rates shown inFig. 5c

	Estimate	Std. error	df	t value	p value
Intercept	6.31	4.78	6355	1.320	.187
Initial Set Complexity	9.21	6.02	8625	1.530	.126
Change in Complexity	-5.07	8.08	3731	-6.279	.000 ***
Sum of shared features	3.22	8.03	8351	4.006	.000 ***
Initial Set Complexity × Change in Complexity	4.96	0.94	3625	5.550	.000 ***
Initial Set Complexity × Sum of shared features	-5.86	1.33	4788	-4.407	.000 ***

the blocks (while still remaining higher than in the other conditions). This decreasing pattern, which did not appear for other conditions and seemed to be specific to the condition 4-2-8, may indicate that participants were mostly deceived by this condition at the beginning of the task, but could adapt progressively to the task at hand to encode features better.

#### Comparison between Experiment 1 and Experiment 2

Finally, we wanted to verify whether memory performance between Experiment 1 and Experiment 2 yielded comparable results for variables that were similar in the two



**Fig. 6** False alarm rates for conditions with zero new features and where complexity decreased with the lure between the first blocks and the other blocks. Error bars represent  $\pm 1$  SE. *Note.* The triplets of numbers used to code conditions correspond to, respectively, Initial Set Complexity, Final set complexity, and Number of shared features between the initial set and the probe

experiments. To test the robustness of our results the simplest way, we only focused on the single effect of the factor Initial set complexity. The reason is that the two experiments used the same levels of Initial Set Complexity (i.e., 4, 6, and 10). The two other factors were globally not comparable and were left out (simply because the changes in complexity that were manipulated were not necessarily the same between the two experiments). Because we could not systematically focus on comparable conditions between the two experiments that involved a change in complexity, the present analysis only focused on hits, for which there was no change.

To analyze the effect of the factor Initial Set Complexity on hits, the hit rates and hit RTs were each converted to z scores. The z scores were computed by aggregating performance per participant and transforming and computing their z scores based on the mean and SD values computed for each experiment. This transformation was thought to reduce cohort effects between experiments. As can be seen in Fig. 7, we observed better performance with lower complexity in both experiments, and memory performance was very close between the two experiments.

For both dependent variables, the Bayesian repeatedmeasures ANOVA suggests that the model that only included the factor Initial Complexity showed evidence against the null hypothesis (BF<sub>10</sub> > 1 e+10). This model was also considered better than the null model that did not include the factor Initial Complexity, both dependent variables (BFm > 1 e+14). However, for both variables, including the factor Experiment or the interaction term did not increase the model probability (BFm < 3). See Table 7. We can thus conclude that the effect of complexity was quite robust across experiments (i.e., the effect was constant under a variety of conditions).

# Discussion

In agreement with our findings in Experiment 1, the results of Experiment 2 indicate an effect of stimulus set complexity with higher Initial Set Complexity corresponding to lower memory performance, and this effect cannot be solely explained by similarity effects (measured by the number of shared features between the probe and the stimuli). For the factor Change of Complexity (i.e., the change of complexity hypothesized to occur considering that the probe mixes with the stimuli), the results showed variable effects depending on Initial Set Complexity like in Experiment 1. When initial complexity is low, a large decrease in FAs following an increase in complexity with the probe can be observed. This corresponds to the greater ability of the participants to detect a change when the stimuli can be more easily encoded. However, with moderate initial complexity this decrease in FAs is less pronounced, and no decrease is detectable when the stimulus set complexity is initially high. Overall, because the rate of FAs is comparably low when complexity increase with introduction of the lure, we conclude that the FAs are essentially provoked by a decrease in complexity, which was predicted by the overcompression hypothesis. Participants seem to encode stimuli with lossy representations, which eventually lure them when a lure that does not modify the complexity level much is introduced. We also observed overall lower performance in this case, indicating that the decision was



Fig. 7 Effects of Initial Set Complexity in Experiment 1 and Experiment 2 measured by (a) hit rates converted to Z scores, and (b) hit response times (RTs) converted to Z scores. Error bars represent  $\pm 1$  SE

 Table 7
 Model comparison for the factors Initial Set Complexity and Experiment based on (a) hit rates converted to Z scores, and (b) hit response times (RTs) converted to Z scores shown in Fig. 7

Model comparison							
Models	P(M)	P(M data)	BF <sub>M</sub>	BF 10	error %		
a)							
Null model (incl. subject)	0.200	3.945 <sup>e</sup> -16	1.578 <sup>e</sup> -15	1			
Initial Set Complexity	0.200	0.722	10.375	1.829 <sup>e+</sup> 15	0.548		
Initial Set Complexity + Experiment	0.200	0.253	1.357	6.420 <sup>e+</sup> 14	3.179		
Initial Set Complexity + Experiment + Initial Set Complexity * Experiment	0.200	0.025	0.102	6.327 <sup>e+</sup> 13	4.739		
Experiment	0.200	1.065 <sup>e</sup> -16	4.259 <sup>e</sup> -16	0.270	0.584		
<b>b</b> )							
Null model (incl. subject)	0.200	1.662 <sup>e</sup> -15	6.648 <sup>e</sup> -15	1			
Initial Set Complexity	0.200	0.567	5.238	3.412 <sup>e+</sup> 14	1.343		
Initial Set Complexity + Experiment	0.200	0.389	2.545	2.340 <sup>e+</sup> 14	5.222		
Initial Set Complexity + Experiment + Initial Set Complexity * Experiment	0.200	0.044	0.185	2.655 <sup>e+</sup> 13	7.866		
Experiment	0.200	9.681 <sup>e</sup> -16	3.873 <sup>e</sup> -15	0.583	3.662		

difficult even though the decision did not result in FAs. These results are in accordance with Experiment 1, which only showed an effect of change in complexity with a moderate complexity of the initial set. Thus, Experiment 2 extends the results of Experiment 1 by indicating a clearer interaction between Initial Stimulus Set Complexity and Change of Complexity produced by a lure.

# General discussion

Previous studies have not yet considered the hypothesis that a cognitive compression process could also lead to over-simplified information due to limitations of the compression process itself. We thus hypothesized that the overregularization of features of a visual scene could produce false recognition of patterns, not because of storage capacity limits but because of compression limits. In two experiments, we prompted a compression process by using a material for which the underlying information varied in compressibility, and our analysis targeted how much compression could take place successfully for a given complex stimulus. In Experiment 1, we used a diverse set of patterns allowing a great number of haphazard variations of features to study different associations of stimulus sets and probes, whereas Experiment 2 targeted a better-controlled set of stimuli to separate the complexity of the visual scene from the complexity of the probe.

Our findings summarized in Table 8 confirm previous results that working memory performance is higher when

regular patterns are present in visual material, confirming that data compression can occur on the spot for newly encountered visual material (Chekaf et al., 2016). This result can help understand why chunking processes seem to reduce the load in working memory (Cowan et al., 2012; Norris & Kalm, 2018; Thalmann et al., 2019). A potential role of working memory might be to compress information to form chunks in long-term memory, which would allow freeing up capacity for subsequent encoded material. The question for future research is whether a chunk is initially encoded into WM or if long-term memory is sufficiently autonomous to compact diverse elements to form progressively a compact chunk. We also observed higher performance for more compressible sets, which can be attributed to the idea that shorter compressed representations take less time to be decompressed. This was observed in a previous research that showed that the RTs taken to categorize learned stimuli depend on their compressibility (Bradmetz & Mathy, 2008), or based on simple logical rules and exception looking like a compressibility metric (see Fific et al., 2010; Nosofsky et al., 1994). Our results thus confirm that perceptual grouping can help observers summarize information to take decisions that can be faster when data compression is achieved.

We believe that our results bring additional outcomes. Our results show that by attempting to benefit from redundant information in a display set, participants can make errors typical of a compression process. False recognition could be due to interference when the display set and the target shared similar features (Oberauer & Lin,

DV	Effect of	Exp.	Significant effect	Main result	Benefit or drawback of compression
Drift Rates	Initial Set Complexity	1	NO main effect, but sig. paired comparisons and an interaction with Change in Complexity	Decrease with increase of Initial Set Complexity	Benefit
		2	YES and an interaction with Change in Complexity	Decrease with increase of Initial Set Complexity	Benefit
	Change in Complexity	1	YES and sig. paired comparisons and one interaction with Initial Set Complexity	Increase when Change in Complexity increases	Drawback
		2	YES and one interaction with Initial Set Complexity	Increase when Change in Complexity increases for low or medium Initial Set Complexity	Drawback
FA Rates	Initial Set Complexity	1	NO main effect, but sig. paired comparisons and an interaction with Change in Complexity	Increase when Initial Set Complexity increases When Initial Complexity is low, Change in Complexity has an effect	Benefit and Drawback
		2	NO main effect, but sig. paired comparisons and an interaction with Change in Complexity	Increase when Initial Set Complexity increases When Initial Complexity is low, Change in Complexity has an effect	Benefit and Drawback
	Change in Complexity	1	NO main effect, But one sig paired comparison: 6-6-4 vs 6-10-4	Decrease when the lure increases complexity	Drawback
		2	YES and one interaction with Initial Set Complexity	Decrease when Change in Complexity increases for low or medium Initial Set Complexity	Drawback

Table 8 Summary of main results for each dependant variable (DV; i.e., False alarm rate (FA rates) and Drift rates) across Experiments

Each row indicates if the result suggests a benefit of compression (a facilitation of memorization with a lower complexity) or a drawback (a typical error of recognition when the lure does not increase complexity)

2017), but this aspect was controlled. We can only infer from our observation that different patterns of errors occurred with different experimental conditions based on the compressibility metric that participants tended to overcompress information, which caused a false recognition of a lure, in particular when introduction of the probe decreased the complexity of the initial set (the design here simulating extra compression of information). These findings can contribute to better understanding why some memory errors seem less costly than others (Sims, 2015). Effectively, there are two ways of benefiting from structured information using data compression: participants can encode information based on a lossless compression process (i.e., the initial object is faithfully represented), or participants attempt to compress information more maximally, which can results in a lossy compressed representations (i.e., the initial object might not be faithfully compressed; see Haladjian and Mathy (2015), who studied memory precision for spatial information).

As in Mewhort and Johns (2000) we observed an effect of the lure that contained features not presented on the display set on the FA rate in both experiments. However, when the lure was entirely made of features presented on the display set, we still observed an effect of Initial Set Complexity (Exp. 1 and Exp. 2) and an effect of interaction between Initial Set Complexity and Change in Complexity in Experiment 2. This interaction was not observed in Experiment 1 because these conditions where the most complex, therefore limiting the effect of Change in Complexity.

From a theoretical stand point, the hypothetical slots in working memory could integrate chunks formed by a lossy compression process, which would explain why some data are in favor of both discrete capacity limit memory models and models based on shared resource, which predict variation in resolution in working memory. Lossy compression could also relate to effects of memory distortion that have been found in studies in which ensemble statistics can bias memory for individual items (Brady & Alvarez, 2011). Our results seem to indicate that a lossy compression can occur on the spot, which confirms a study by Nassar et al. (2018), who showed that participants can implement a lossy form of data compression to improve working memory performance with reinforcement learning. Following Nassar et al., we believe that lossy compression effects could help reconcile theories of working memory capacity based on either discrete limitation (i.e., a series of slots) or continuous limitations (i.e., a divisible resource). Compression in our study took place based on conjunction of features, instead of simply being driven by the mean of different values of a single continuous dimension as in Brady and Alvarez (2011). Our data, however, contrast with those of Nassar et al. because we observed in an additional analysis that performance (a greater number of hits and a lower number of FAs) significantly increased as a function of trial number (additional mixed models showed a significant effect of trial number on these two dependent variables, with both p < .0003). In our case, this can mean that the sum of information in our material was adequate to fit the processing ability of participants, who could progressively adopt a lossless compression strategy rather than a lossy compression strategy. In sum, our particular material might have induced specific compression processes by letting participants adapt to the task at hand, and the specific range of complexity in our material could also explain why our findings diverge from those of Nassar.

Our main conclusion is that memory performance for visually presented material depends on compressibility. For each visual display, memory was successful to the degree that the set of stimuli could be faithfully compressed, and memory seemed to be unsuccessful for the inverse reason that some of the most compressible sets of stimuli could not be faithfully compressed and appeared to be over-compressed. Although we used a metric based on the logical structure of information, it does not mean that we consider that compression exclusively operates by converting visual information into logical descriptions. We do not assume that stimuli are encoded into a verbal format or are strictly represented qua Boolean functions, rather than into a visual format. There seems to be an ease with which a display can be stored depending on the internal structure of information that specifies how features are distributed among objects, thus allowing a certain degree of compression. Encoding of the features could be either verbal or visual information stored more minimally, because both types of information were likely to be encoded (as we intentionally used a procedure letting participants encode the stimuli freely). In what exact format participants encoded information remains uncertain and needs further study. Unfortunately, it is difficult to consider that the subjective reports can provide clues to which modality was preferred to encode information, as introspection can only reveal partial aspects of the encoding process, not mentioning that introspection can be misleading (e.g., the participant can have the impression that verbalizing the features helped them whereas it is mostly how visual information was implicitly encoded that drove memory performance) and maybe their verbalization interferes with this process. Further studies should therefore more simply consider using stimuli followed by masks and articulatory suppression during the process.

To conclude, our study shows that data compression can occur in the short-term memorization of stimulus sets made of discrete features. This phenomenon can be detected in performance correlated with complexity and by studying typical memory errors that seem to derive from a lossy compression process.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.3758/s13414-021-02333-x.

Authors' contributions The authors' contributions are reported following the CRediT taxonomy (see https://casrai.org/credit/). Conceptualization, N.C., L.L., F. L, and F.M.; Methodology, L.L. and C.A.; Formal Analysis, L.L..; Writing, N.C., L.L., and F.M. Funding Acquisition (F. M.).

**Funding** This research was supported in part by a grant from the Agence Nationale de la Recherche (ANR-17-CE28-0013-01) awarded to Fabien Mathy.

**Data availability** The data for all experiments are available in the OSF repository at https://osf.io/ehjrw/?view\_only=0373989a406643d0ad0a8548e4829ff3

#### Declarations

**Ethics approval** The experiment was approved by the local ethics committee (CERNI) of the Université Côte d'Azur.

**Consent to participate** The experiment was conducted with the informed consent of the participants.

**Consent for publication** Participants signed informed consent regarding publishing their data.

**Conflicts of interest/Competing interests** The authors have no conflicts of interest to declare that are relevant to the content of this article.

# Appendix

Example of display and probe for each condition of Experiment 1 and Experiment 2







# References

- Akaike, H. (1987). Factor analysis and AIC. In Selected papers of hirotugu akaike (p. 371-386). Springer.
- 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(2), 106-111.
- Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological science*, 18(7), 622-628.
- Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of* vision, 9(10), 7-7.
- Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, 321(5890), 851-854.
- Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia*, 49(6), 1622-1631.
- Bradmetz, J., & Mathy, F. (2008). Response times seen as decompression times in Boolean concept use. *Psychological Research*, 72(2), 211-234.
- Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory : Ensemble statistics bias memory for individual items. *Psychological science*, 22(3), 384-392.
- Brady, T. F., & Alvarez, G. A. (2015). No evidence for a fixed object limit in working memory : Spatial ensemble representations inflate estimates of working memory capacity for complex objects. *Journal* of Experimental Psychology: Learning, Memory, and Cognition, 41(3), 921.
- 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(4), 487-502. https://doi.org/10.1037/a0016797

- 2863
- Brady, T. F., & Tenenbaum, J. B. (2013). A probabilistic model of visual working memory : Incorporating higher order regularities into working memory capacity estimates. *Psychological Review*, 120(1), 85.
- Brown, V. A. (2020). An approachable introduction to linear mixed effects modeling with implementation in R.
- Chekaf, M., Cowan, N., & Mathy, F. (2016). Chunk formation in immediate memory and how it relates to data compression. *Cognition*, 155, 96-107.
- Chekaf, M., Gauvrit, N., Guida, A., & Mathy, F. (2018). Compression in working memory and its relationship with fluid intelligence. *Cognitive Science*, 42, 904-922.
- Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck : A fundamental constraint on language. *Behavioral and Brain Sciences*, 39.
- Cowan, N. (2001). The magical number 4 in short-term memory : A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87-114.
- Cowan, N., Blume, C. L., & Saults, J. S. (2013). Attention to attributes and objects in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 731.
- Cowan, N., Rouder, J. N., Blume, C. L., & Saults, J. S. (2012). Models of verbal working memory capacity : What does it take to make them work? *Psychological review*, 119(3), 480.
- De Lillo, C. (2004). Imposing structure on a Corsi-type task : Evidence for hierarchical organisation based on spatial proximity in serialspatial memory. *Brain and Cognition*, 55(3), 415-426.
- Dry, M., Preiss, K., & Wagemans, J. (2012). Clustering, randomness, and regularity : Spatial distributions and human performance on the traveling salesperson problem and minimum spanning tree problem. *The Journal of Problem Solving*, 4(1), 2.
- Feldman, J. (1999). The role of objects in perceptual grouping. Acta Psychologica, 102(2-3), 137-163.
- Feldman, J. (2000). Minimization of Boolean complexity in human concept learning. *Nature*, 407(6804), 630-633.
- Feldman, J. (2003). A catalog of Boolean concepts. *Journal of Mathematical Psychology*, 47(1), 75-89.
- Ferrer-i-Cancho, R., Hernández-Fernández, A., Lusseau, D., Agoramoorthy, G., Hsu, M. J., & Semple, S. (2013). Compression as a universal principle of animal behavior. *Cognitive Science*, 37(8), 1565-1578.
- Fific, M., Little, D. R., & Nosofsky, R. M. (2010). Logical-rule models of classification response times : A synthesis of mental-architecture, random-walk, and decision-bound approaches. *Psychological Review*, 117(2), 309.
- Fougnie, D., & Alvarez, G. A. (2011). Object features fail independently in visual working memory : Evidence for a probabilistic featurestore model. *Journal of vision*, 11(12), 3-3.
- Gao, Z., Gao, Q., Tang, N., Shui, R., & Shen, M. (2016). Organization principles in visual working memory : Evidence from sequential stimulus display. *Cognition*, 146, 277-288.
- Gauvrit, N., Singmann, H., Soler-Toscano, F., & Zenil, H. (2016). Algorithmic complexity for psychology : A user-friendly implementation of the coding theorem method. *Behavior research methods*, 48(1), 314-329.
- Haladjian, H. H., & Mathy, F. (2015). A snapshot is all it takes to encode object locations into spatial memory. *Vision research*, 107, 133-145.
- Hardman, K. O., & Cowan, N. (2015). Remembering complex objects in visual working memory : Do capacity limits restrict objects or features? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(2), 325.
- Hutter, M. (2004). Universal artificial intelligence : Sequential decisions based on algorithmic probability. Springer Science & Business Media.
- Jiang, Y., Chun, M. M., & Olson, I. R. (2004). Perceptual grouping in change detection. *Perception & Psychophysics*, 66(3), 446-453.

- Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 683.
- Kemp, C. (2012). Exploring the conceptual universe. *Psychological Review*, 119(4), 685.
- Kibbe, M. M., & Kowler, E. (2011). Visual search for category sets : Tradeoffs between exploration and memory. *Journal of Vision*, *11*(3), 14-14.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87-102.
- Korjoukov, I., Jeurissen, D., Kloosterman, N. A., Verhoeven, J. E., Scholte, H. S., & Roelfsema, P. R. (2012). The time course of perceptual grouping in natural scenes. *Psychological Science*, 23(12), 1482-1489.
- Lafond, D., Lacouture, Y., & Mineau, G. (2007). Complexity minimization in rule-based category learning : Revising the catalog of Boolean concepts and evidence for non-minimal rules. *Journal of Mathematical Psychology*, 51(2), 57-74.
- Li, M., & Vitányi, P. (2008). An introduction to Kolmogorov complexity and its applications (Vol. 3). Springer.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279.
- Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, 17(3), 347.
- Mathy, F., & Bradmetz, J. (2004). A theory of the graceful complexification of concepts and their learnability. *Cahiers de Psychologie Cognitive-Current Psychology of Cognition*, 22, 39-80.
- Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition*, 122(3), 346-362.
- Mewhort, D. J. K., & Johns, E. E. (2000). The extralist-feature effect : Evidence against item matching in short-term recognition memory. *Journal of Experimental Psychology: General*, 129(2), 262.
- Morey, C. C., Cong, Y., Zheng, Y., Price, M., & Morey, R. D. (2015). The color-sharing bonus : Roles of perceptual organization and attentive processes in visual working memory. *Archives of Scientific Psychology*, 3(1), 18.
- Nassar, M. R., Helmers, J. C., & Frank, M. J. (2018). Chunking as a rational strategy for lossy data compression in visual working memory. *Psychological Review*, 125(4), 486.
- Ngiam, W. X., Brissenden, J. A., & Awh, E. (2019). "Memory compression" effects in visual working memory are contingent on explicit long-term memory. *Journal of Experimental Psychology: General*, 148(8), 1373.
- Norris, D. G., & Kalm, K. (2018). *Chunking and redintegration in verbal short-term memory.*
- Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plusexception model of classification learning. *Psychological Review*, 101(1), 53.
- Oberauer, K., & Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Memory & Cognition*, 41(8), 1212-1227.
- Oberauer, K., & Lin, H.-Y. (2017). An interference model of visual working memory. *Psychological Review*, 124(1), 21.
- Orbán, G., Fiser, J., Aslin, R. N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences*, 105(7), 2745-2750.
- Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. Journal of Neuroscience Methods, 162(1-2), 8-13.

- Peterson, D. J., & Berryhill, M. E. (2013). The Gestalt principle of similarity benefits visual working memory. *Psychonomic Bulletin & Review*, 20(6), 1282-1289.
- Quinlan, P. T., & Cohen, D. J. (2012). Grouping and binding in visual short-term memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 38(5), 1432.
- Ramzaoui, H., & Mathy, F. (2021). A compressibility account of the color-sharing bonus in working memory. *Attention, Perception, & Psychophysics*, 1-16.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59.
- Reder, L. M., Liu, X. L., Keinath, A., & Popov, V. (2016). Building knowledge requires bricks, not sand : The critical role of familiar constituents in learning. *Psychonomic Bulletin & Review*, 23(1), 271-277.
- Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences*, 105(16), 5975-5979.
- Saiki, J. (2019). Robust color-shape binding representations for multiple objects in visual working memory. *Journal of Experimental Psychology: General.*
- Sargent, J., Dopkins, S., Philbeck, J., & Chichka, D. (2010). Chunking in spatial memory. *Journal of Experimental Psychology: Learning*, *Memory, and Cognition*, 36(3), 576.
- Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. *Nature human behaviour*, 4(11), 1156-1172.
- Sims, C. R. (2015). The cost of misremembering : Inferring the loss function in visual working memory. *Journal of vision*, 15(3), 2-2.
- Singmann, H., & Kellen, D. (2019). An introduction to mixed models for experimental psychology. *New methods in cognitive psychology*, 28, 4-31.
- Soler-Toscano, F., Zenil, H., Delahaye, J.-P., & Gauvrit, N. (2014). Calculating Kolmogorov complexity from the output frequency distributions of small Turing machines. *PloS one*, 9(5), e96223.
- Thalmann, M., Souza, A. S., & Oberauer, K. (2019). How does chunking help working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(1), 37.
- Vigo, R. (2006). A note on the complexity of Boolean concepts. *Journal of Mathematical Psychology*, 50(5), 501-510.
- Wagenmakers, E.-J., Van Der Maas, H. L., & Grasman, R. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, 14(1), 3-22.
- Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, 131(1), 48.
- Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review*, 10(1), 80-87.
- Xu, Y. (2002). Limitations of object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, 28(2), 458.
- Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, 440(7080), 91-95.
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453(7192), 233.

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