Cognition 122 (2012) 346–362

Contents lists available at SciVerse ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

What's magic about magic numbers? Chunking and data compression in short-term memory

Fabien Mathy [a,*], Jacob Feldman [b]

[a] Université de Franche-Comté, 30-32 rue Mégevand, 25030 Besançon Cedex, France
[b] Rutgers University, New Brunswick, USA

ARTICLE INFO

Article history:
Received 1 February 2011
Revised 25 October 2011
Accepted 2 November 2011
Available online 15 December 2011

Keywords:
Chunk
Chunking
Compression
Short-term memory
Capacity
Span

ABSTRACT

Short term memory is famously limited in capacity to Miller's (1956) magic number 7 ± 2—or, in many more recent studies, about 4 ± 1 "chunks" of information. But the definition of "chunk" in this context has never been clear, referring only to a set of items that are treated collectively as a single unit. We propose a new more quantitatively precise conception of chunk derived from the notion of Kolmogorov complexity and compressibility: a chunk is a unit in a *maximally compressed* code. We present a series of experiments in which we manipulated the compressibility of stimulus sequences by introducing sequential patterns of variable length. Our subjects' measured digit span (raw short term memory capacity) consistently depended on the length of the pattern *after compression*, that is, the number of distinct sequences it contained. The true limit appears to be about 3 or 4 distinct chunks, consistent with many modern studies, but also equivalent to about 7 uncompressed items of typical compressibility, consistent with Miller's famous magical number.

1. Introduction

In a famous paper, Miller (1956) proposed that the capacity of short-term memory (STM) is limited to a "magical number" of about seven (plus or minus two) items.[1] This limit is usually expressed in terms of "chunks" (Anderson, Bothell, Lebiere, & Matessa, 1998; Gobet et al., 2001; Simon, 1974; Tulving & Patkau, 1962), meaning groups of items that have been collected together and treated as a single unit, in part to accommodate the observation that apparent span may be increased if items can be readily grouped together into larger units. For example, amid a sequence of letters the familiar string USA or the repeating pattern *BBB* might each serve as a single chunk, rather than as three separate items each. An extreme example of chunking is the subject S.F. discussed in Ericsson, Chase, and Faloon (1980), who despite average intelligence was able to increase his apparent digit span to almost 80 digits by devising a rapid recoding system based on running times, which allowed him to group long sequences of digits into single chunks.

The capacity limit is traditionally attributed to forgetting by rapid time-based decay (Baddeley, 1986; Barouillet, Bernardin, & Camos, 2004; Barouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007; Burgess & Hitch, 1999; Henson, 1998; Jonides et al., 2008; Nairne, 2002; Page & Norris, 1998) or mutual interference between items (Lewandowsky, Duncan, & Brown, 2004; Nairne, 1990; Oberauer & Kliegl, 2006). The span is also substantially influenced by the spoken duration of the constituent items, a result which runs against a constant chunk hypothesis and which has been interpreted in terms of a phonemically-based store of limited temporal capacity (Baddeley, Thomson, & Buchanan, 1975; Burgess & Hitch, 1999; Estes, 1973; Zhang & Simon, 1985). Though verbal STM is well known to depend on phonological encoding (Baddeley, 1986; Chen & Cowan, 2005), the sometimes dramatic influence of chunking points to abstract unitization mechanisms that are still poorly understood.

* Corresponding author.
E-mail address: fabien.mathy@univ-fcomte.fr (F. Mathy).

[1] According to the Science Citation Index (Kintsch & Caciopo, 1994) this paper is the most frequently cited article in the history of *Psychological Review*.

0010-0277/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2011.11.003

Notwithstanding the fame of Miller's number (Baddeley, 1994), many more recent studies have converged on a smaller estimate of STM capacity of about four items (Baddeley & Hitch, 1974; Brady, Konkle, & Alvarez, 2009; Broadbent, 1975; Chase & Simon, 1973; Estes, 1972; Gobet & Clarkson, 2004; Halford, Baker, McCredden, & Bain, 2005; Halford, Wilson, & Phillips, 1998; Luck & Vogel, 1997; Pylyshyn & Storm, 1988, 2008). The concept of working memory (Baddeley, 1986; Engle, 2002) has emerged to account for a smaller "magic number" that Cowan (2001) estimated to be 4 ± 1 on the basis of a wide variety of data.

Broadly speaking, the discrepancy between the two capacity estimates seems to turn on whether the task setting allows chunking (Cowan, 2001). Generally, four is the capacity that has been observed when neither rehearsal nor long-term memory can be used to combine stimulus items (i.e., to chunk), while seven is the limit when chunking is unrestricted. Hence the two limits might be fully reconciled if only chunking were more completely understood.

Yet half a century after Miller's article, the definition of a chunk is still surprisingly tentative. Chunks have been defined as groups of elements (Anderson & Matessa, 1997; Bower & Winzenz, 1969; Cowan, 2010; Cowan, Chen, & Rouder, 2004; Farrell, 2008; Hitch, Burgess, Towse, & Culpin, 1996; Ng & Maybery, 2002; Ryan, 1969; Wickelgren, 1964), but exactly which groups remains unclear unless they result from statistical learning (Perruchet & Pacton, 2006; Servan-Schreiber & Anderson, 1990). Cowan (2001) defines a chunk as "a collection of concepts that have strong associations to one another and much weaker associations to other chunks concurrently in use" and Shiffrin and Nosofsky (1994) as "a pronounceable label that may be cycled within short-term memory". Most attempts to define chunks are somewhat vague, ad hoc, or severely limited in scope, especially when they apply only to verbally encoded material (Shiffrin & Nosofsky, 1994; Stark & Calfee, 1970), making it difficult for them to explain the existence of chunking-like processes in animal learning (Fountain & Benson, 2006; Terrace, 1987, 2001). The current consensus is that (1) the number seven estimates a capacity limit in which chunking has not been eliminated (2) there is a practical difficulty in measuring chunks and how they can be packed and unpacked into their constituents.

In this paper we propose a new conception of chunk formation based on the idea of *data compression*. Any collection of data (such as items to be memorized) can be faithfully represented in a variety of ways, some more compact and parsimonious than others (Baum, 2004; Wolff, 2003). The size of the most compressed (lossless) representation that faithfully represents a particular sequence is a measure of its inherent randomness or complexity, sometimes called its *Kolmogorov complexity* (Kolmogorov, 1965; Li & Vitányi, 1997).

Simpler or more regular sets can be represented more compactly by an encoding system that takes advantage of their regularities, e.g. repetitions and symmetries. As an upper bound, a maximally complex sequence of N items will require about N slots to encode it, while at the other extreme an extremely repetitive string may be compressed into a form that is much smaller than the original string. Incompressibility as a definition of subjective randomness has some empirical support (Nickerson, 2002). Kolmogorov complexity has a number of cognitive correlates (Chater & Vitányi, 2003); for example simpler categories are systematically easier to learn (Feldman, 2000; Pothos & Chater, 2002).

In this paper, we ask whether complexity influences the ease with which material can be committed to short-term memory. Our hypothesis, that simpler material is more easily memorized, follows directly from the fact that—by definition—complexity determines the size of a maximally compressed representation.

If so, the true limits on capacity depend on the size of this compressed code, leading to our view that a "chunk" is really a unit in a maximally compressed code. The following experiments test this hypothesis by systematically

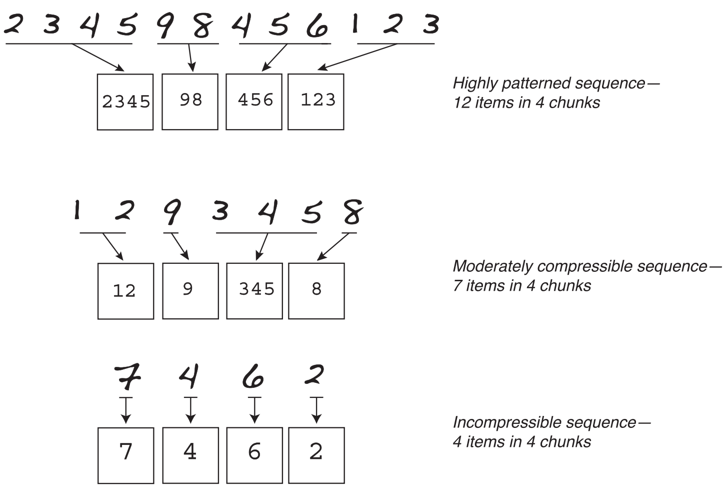

Fig. 1. The number of items that can be compressed into four "chunks" depends on the complexity of these material. Completely incompressible (maximum Kolmogorov complexity) sequences (bottom) require one chunk per item. Sequences of moderate complexity (middle) might allow 7 items to be compressed into 4 chunks, leading to an apparent digit span of 7. Highly patterned (regular) sequences might (top) allow even larger numbers of items to be compressed into the same four slots.