Forming Concepts of Mozart and Homer Using Short-Term and Long-Term Memory: A Computational Model Based on Chunking

Dmitry Bennett (D.Bennett@liverpool.ac.uk)
Department of Psychological Sciences, Bedford Street South,
University of Liverpool, L69 7ZA, United Kingdom

Fernand Gobet (F.Gobet@lse.ac.uk)
Centre for Philosophy of Natural and Social Science, Houghton Street,
London School of Economics, WC2A 2AE, United Kingdom

Peter Lane (p.c.lane@herts.ac.uk)
Department of Computer Science, College Lane,
University of Hertfordshire, AL10 9AB, United Kingdom

Abstract

A fundamental issue in cognitive science concerns the mental processes that underlie the formation and retrieval of concepts in the short-term and long-term memory (STM and LTM respectively). This study advances Chunking Theory and its computational embodiment CHREST to propose a single model that accounts for significant aspects of concept formation in the domains of literature and music. The proposed model inherits CHREST's architecture with its integrated STM/LTM stores, while also adding a moving attention window and an "LTM chunk activation" mechanism. These additions address the overly destructive nature of primacy effect in discrimination network based architectures and expand Chunking Theory to account for learning, retrieval and categorisation of complex sequential symbolic patterns – like real-life text and written music scores. The model was trained through exposure to labelled stimuli and learned to categorise classical poets/writers and composers. The model categorised previously unseen literature pieces by Homer, Chaucer, Shakespeare, Walter Scott, Dickens and Joyce, as well as unseen sheet music scores by Bach, Mozart, Beethoven and Chopin. These findings offer further support to mechanisms proposed by Chunking Theory and expand it into the psychology of music.

Keywords: categorisation; CHREST; concept; chunking; learning; literature; long-term memory; music; short-term memory.

Introduction

How do we develop a feeling that a poem sounds "Shakespearian" or that a music piece is "Mozart like"? How do we form, update and apply concepts of Homer and Bach? Indeed, what *are* concepts?

One definition is that concepts are "mental representations of classes of things", with "classes of things" themselves being categories (Murphy, 2002, p.5). According to the naive realist perspective, all concepts have defining features, which, when fulfilled, are sufficient for class membership (Hull, 1920). That view has been convincingly challenged by Wittgenstein (1953) who demonstrated that, while defining features may be necessary (e.g. being animate and breathing for cats), they may not be sufficient.

Prototype approach was put forward to account for concepts' "fuzziness" and suggested that each concept has a prototype with a summary description where the greatest family resemblances point to the most prototypical members (Rosch, 1975). For example, "fruit" might include typical attributes such as having seeds, being edible and growing above the ground – with these attributes being characteristic, rather than defining ones (Hampton, 1979). However, Murphy (2002) notes that prototype theory is vague when it comes to the definition of prototype and, more importantly, lacks clarity on how feature lists are to be determined.

Further, Barsalou (2009) makes the important point that concepts have little meaning when isolated and thus need to be part of an interconnected conceptual web (Goldstone & Steyvers, 2001) while also being linked to sensory-motor processes. "Conceptual representations are modal, not amodal. The same type of representation underlies perception and conception. When the conceptual system represents an object's visual properties, it uses the representations in the visual system; when it represents the actions performed on an object, it uses motor representations" (Barsalou, 2003, p. 521).

The exemplar approach to concept formation (also known as Generalized Concept Model or GCM) has been put forward by Nosofsky (2011) to address some of the concerns above. According to the exemplar theory, instead of operating on abstract lists of features, the memory system stores large numbers of specific instances; a typicality gradient may thus be derived from the underlying pattern. For example, it would derive the "bird" concept from a distribution incorporating commonly encountered pigeons as well as a rarely seen penguin. One other strong point of the exemplar model of classification is that it has quantitative

analysis at its core, addressing the inescapable ambiguities that necessarily accompany verbal theories.

The current study aims to address several criticisms that may be aimed at the research above. Firstly, when applied to real-world natural category domains, the exemplar approach's typicality gradient is devised by computing probabilities based on human participants' (or, indeed, experts') responses within some narrow modality. For example, a psychological model of categorisation of rocks (Nosofsky, Sanders, & McDaniel, 2018) was made possible by participants providing similarity judgments among the pairs of rock specimens and/or by deriving corresponding dimensions from geology textbooks. Student participants have then scored various rock stimuli along these very dimensions with the scores being input into GCM. For instance, a participant may score a piece of granite as "lightness/darkness of colour = 4; average grain size = 2; shininess = 8; roughness/smoothness = 2"; these values may then be put into GCM equations to model recognition of granite and continue to train it with more participants and more rocks. One implication here is that the GCM account of concept formation relies on the conceptions of particular experts (geology in this case) to provide it with the required dimensions and, outside of the imposed similarity metric, does not allow for participants to develop their own subjective dimensions (e.g. "these stones look like the ones that made up my grandmother's fireplace" dimension). Secondly, GCM offers little insight into how the participants/experts came up with their dimensions and scores in the first place. Thirdly, a fully developed and fully trained model of rock categorisation would once again need expert/human input of dimension data to classify poems (novels, music pieces… ad infinitum) – it cannot account for learning concepts from raw data.

The other major issue of concept formation theories outlined above is that, outside of vague verbal theorising, they often have little to say on how the formation of concepts is rooted in fundamental psychological mechanisms such as STM, LTM and the perceptual apparatus. For example, it is difficult to see how the "magic number 7 plus or minus 2" (Miller, 1956) is related to the formation, update and retrieval of a stored concept of a rock (a poem, a music piece) in the exemplar or the prototype theory.

The Outline of Chunking Theory / CHREST Architecture

Before proceeding with what the current model has to offer to the discussion above, we will briefly introduce the theory and the modelling architecture itself.

The origins of Chunking Theory can be traced to 1959 (Chase & Simon, 1973; Simon, 1974). Since then, the theory has captured at least 20 findings in verbal learning research (Richman, Simon, & Feigenbaum, 2002), shed light on STM and distilled seminal work on perception and memory in chess (Chase & Simon, 1973; De Groot, 1965; Gobet & Simon, 1996c). All of the research above (plus more) is encapsulated in computational architectures, first EPAM *Elementary Perceiver and Memorizer* (Richman, Staszewski, & Simon, 1995), and now CHREST *Chunking Hierarchy and Retrieval Structures* (Gobet & Lane, 2012), which embodies most of EPAM's mechanisms.

As its name suggests, Chunking Theory is founded on the proposed mechanism of *chunking* (Simon, 1974). While a *chunk* can be defined as a meaningful unit of information constructed from elements that have strong associations between each other (e.g. several digits making up a telephone number or a group of letters and digits making up a postal address), *chunking* is the process of creating and updating chunks in the cognitive system. Although chunks themselves vary between people due to personal differences, the chunking *mechanism* is largely invariant across domains, individuals and cultures (Chase & Simon, 1973; De Groot, 1965; Gobet et al., 2001; Miller, 1956).

CHREST is a self-organising computer model that simulates human learning processes. The patterns that are processed by CHREST are *symbolic* – they are meaningful and are represented in identical ways for objects inside (cognition) and outside (input) the architecture. Thus, CHREST is an example of a symbolic cognitive architecture (often contrasted with subsymbolic connectionist neural network approach, although the two are similar in their focus on perception as the primary attribute of cognition). Patterns are assumed to be composite objects made up of primitives: for example, a collection of letters making up words, a collection of words making up sentences, a collection of chords making up musical measures and a collection of measures making up a musical piece.

For both Chunking Theory and CHREST, learning implies incremental growth of an LTM network, a process influenced both by the environmental stimuli – such as literature and music – and by the knowledge that has already been stored (Gobet & Lane, 2012). Concretely, CHREST's STM forms associative links between chunks in its "magic number long" queue-like structure (Miller, 1956). The LTM, on the other hand, consists of a pool of chunks, storage of associations among chunks and memory retrieval structure in the form of

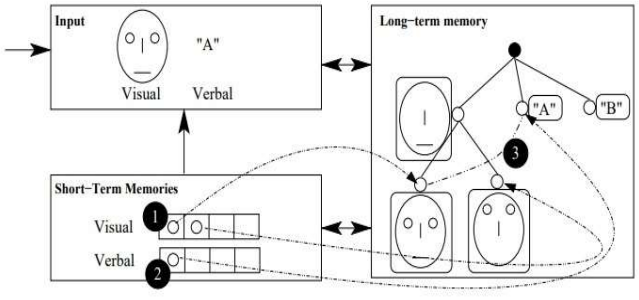

Figure 1. CHREST's learning mechanisms in action: (1) the visual stimulus is sorted through LTM and a pointer to the node retrieved is formed in visual STM; (2) the verbal stimulus is sorted through LTM and pointer to the retrieved node is created in STM; and (3) when a visual pattern and a verbal pattern are stored in STM concurrently, the chunks they elicit are connected together in the LTM.

indexed discrimination network. The latter mechanism is vital for updating existing memory chunks and for integrating perception, STM and LTM into a single coherent whole.

Prior Developments and the Present Study

Historically, modelling concept formation in CHREST has involved two stages: learning a concept by sorting a stimulus through its discrimination net and assigning a lateral "naming link" within the discrimination net to a label node.

The canonical "five-four" categorisation experiment (Smith & Minda, 2000) provides for a simple "toy" demonstration of CHREST's mechanism for forming concepts and differentiating between them. The basic experiment structure is presented in Figure 1. The task here is for an agent to classify a stimulus as either a "type A" or a "type B" face. A stimulus face possesses four binary features, with different interpretations of each face being created by changing the features. The four binary values of the toy faces (A0 – eye height, A1 – eye separation, A2 – nose length, A3 – mouth height) provide 16 different faces. Examples of category A face are typically closer to having all four features turned on, while instances of category B face tend to have the four binary features turned off.

When shown the stimuli, CHREST forms a hierarchy of visual chunks in LTM that contain the visual features of the "faces". At the same time, the same learning process forms nodes with verbal chunks of "A" and "B" labels. Concretely, learning in both domains comes about as the result of revising the LTM discrimination network through *creating* new chunks and *updating* the old chunks with new information. One example of the former process would be the creation of a new "face" representation node with a "large eye height" attribute – if there was no chunk with such a face in LTM at that moment. An illustration of a chunk *update* would be adding the "low mouth" feature to the previously incomplete facial representation of the "low eyes plus low nose" type. When chunks from visual "face" and verbal "label" modalities occupy the same spot in the respective STM queues, a naming link is formed and stored into LTM. Lastly, it should also be noted that the terms "visual" and "verbal" chunks are mere naming conventions for LTM hierarchies that represent distinct domains; the underlying symbolic nature of the patterns and the mechanisms that operate on them are exactly the same in all cases.

The Present Study The present study intended to develop CHREST account of concept formation in non-toy, ecologically valid domains – literature and music. Not only do these domains possess real-life "fuzziness", but they also rely on time-step sensitive sequential data. To achieve this aim, firstly, we tested the present CHREST model stability with sequential non-binary "toy" data. One example of a toy test for resistance to pattern occlusion involved categorisation of city names: "Liverpool = type A", "Manchester = type B". The examples of occluded patterns included "Liverpooz", "Lizerzool", "zzzzLzverzool". Due to CHREST failing to categorise patterns when the occluder is preceding the pattern (e.g. "zLiverpool", "zzzZiverpool" and so on), we then incorporated two new mechanisms: a "moving attention window" and an "LTM chunk activation" measure (more on this below). We trained the model on unabridged works by various authors and composers. We tested categorisation on previously unseen pieces produced by the same authors and composers.

Method

Training and Testing

The training data for the literature categorisation experiment was as follows. For Shakespeare, CHREST was given 140 *Sonnets*, *Romeo and Juliet* and an excerpt of *Midsummer Night's Dream*. The Homer (translated by Samuel Butler) training set had the first four chapters of *The Iliad.* For Dickens, a large excerpt from *David Copperfield* was used.

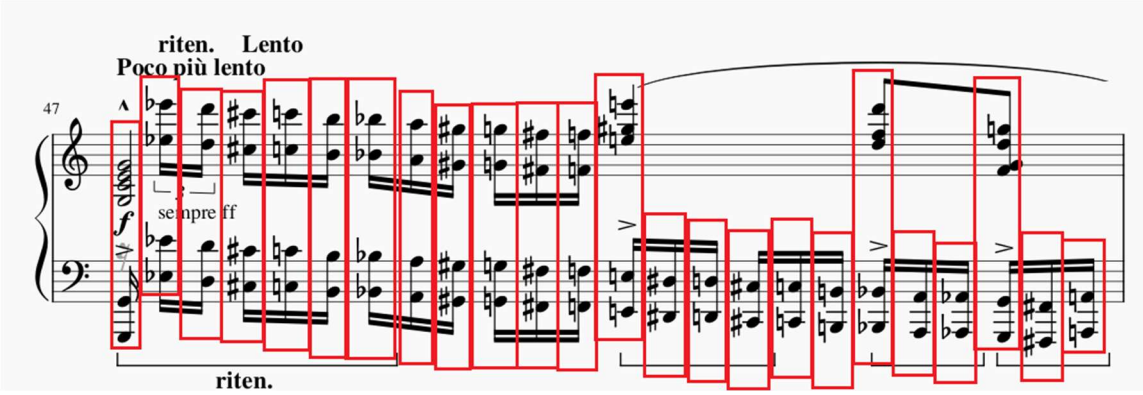

Figure 2. An excerpt from Chopin's Op.48 No1. From the perspective of CHREST, each vertical frame represents a single pattern primitive, equivalent to a word in a text sentence.

For Chaucer, it was *Troilus and Criseyde*. For Walter Scott, the "reading set" included excerpts from *Ivanhoe* and *Rob Roy*. Lastly, the Joyce sample contained the first 4 chapters of *Ulysses*. For every author, there was 300Kb of text in total.

The Shakespeare tests included the remaining *Sonnets*, *The Passionate Pilgrim*, *Venus and Adonis*, *Pericles Prince of Tyre*. The Homer tests contained chapters from *The Odyssey* and the ending chapter of *The Iliad*. The Chaucer tests

included excerpts from *Canterbury Tales*, *Book of Duchesse* and *The Parliament of Fowles*. The Dickens test was comprised of excerpts from *Oliver Twist, The Pickwick Papers*, *A Christmas Carol* and *Tale of Two Cities*. The Walter Scott test category had chapters from *The Black Dwarf*, *Marmion* and *Talisman*. The final chapter of *Ulysses* and excerpts from *Finnegans Wake* and *A Portrait of the Artist as a Young Man* formed the Joyce test.

For music, training included 63 pieces from Bach's *Well-Tempered Clavier (WTC)*; Mozart's *Piano Sonatas No.1-6* and *No.8-12*; Beethoven's *Piano Sonatas No.1-7*; and 14 *Etudes* and 18 *Nocturnes* by Chopin. As with literature, each training category contained approximately 300Kb of text, this time generated from MIDI files.

The music test dataset contained 5 Bach *WTC* pieces; 5 *Sonata* pieces by Mozart; 5 *Sonatas* excerpts for Beethoven; a *Waltz*, *Ballad Op.23*, a *Nocturne* and two *Etudes* for Chopin (for full details see Table 1).

All pieces were transposed to C-major/A-minor key. Full chord complexity and polyphony was preserved, but timings were not kept in the text conversions.

The pattern primitives for the text modality were chosen to be words. In the music case, the primitives were note/chord structures that occupied one time step in a given sequence (see Figure 2).

Following the conclusions of Gobet and Lane (2012), the training samples were split into 20-word phrases (for text) and 2 measures (for music) to avoid forming overly large chunks. The order of the training samples was randomised. No words/notes were removed from either training or testing texts/music scores.

Procedure

Due to CHREST failing to categorise sequential patterns when an occluder is preceding the pattern (e.g. "zLiverpool", "zzzLiverpool" and so on), two important developments to the existing CHREST architecture were proposed.

Firstly, a recursive "sliding attention window" was added to represent the limited scope of a human reader and to allow multiple passes over patterns deemed unfamiliar by the LTM. If a vector of patterns can be represented as input $\boldsymbol{p} = [p_1, p_2, \ldots p_n]$, then CHREST attention window $\boldsymbol{w}$ would fetch $[p_{1+t}, p_{2+t}, \ldots p_{m+t}]$, with m being the span of the window and t being the time step size. Before proceeding to the next step, the window would progressively shrink and fetch sequences $[p_{2+t}, \ldots p_{m+t}]$, $[p_{3+t}, \ldots p_{m+t}]$, $\ldots$ $[p_{m+t-1}, p_{m+t}]$.

Secondly, CHREST now records chunk "activation" – the largest chunk met so far – as a function of an input pattern, to allow for conflict resolution between chunks "voting" for different categories. In the "zzzLizerzool" example, CHREST would iteratively scan the pattern and attempt to retrieve corresponding chunks from its LTM, eventually reading off the naming link/verbal chunk ("type A") from the largest of the retrieved visual chunks ("Liverpool").

The retrieved verbal and visual LTM chunks send pointers to STM, which then enables naming links to be stored in the LTM. It is important to stress that the length of the STM queue is measured in chunks and not in primitives – as we can briefly memorise 7(+/-2) individual digits, but also 7(+/-2) previously learnt telephone numbers (Miller, 1956).

The internal parameters of the model were fixed for the entire duration of the experiment. In particular, the STM size was set to 5 chunks; the maximum size of the attention window was set to 20 words or 2 measures; the likelihood of forming a chunk was set to 1; the time needed to create a new chunk was set to 10 seconds; and the time needed to update a chunk was set to 2 seconds. Music and literature patterns were assigned to the visual modality, while author/composer names were assigned to the verbal modality (with the caveat that this distinction is a mere convention – as discussed above).

If there are m categories, the vector of category labels is $\boldsymbol{c} = [c_1, c_2, \ldots c_m]$, the vector of category specific chunk activations is $\boldsymbol{a} = [a_1, a_2, \ldots a_m]$ and the confidence of a prediction that a pattern belongs to category c_i would be calculated using the equation

$$C(c_i|x) = a_i \,/ \sum_{k=1}^{m} (a_k)$$

where $C(c_i|x)$ is confidence that category label is c_i, given a book or music score x; a_i is the LTM chunks' activation corresponding to that category, and the summation part being the sum of chunk activations across all m categories.

The final important point is that the model will be simultaneously trained on *both* literature and music: its STM/LTM will seamlessly form, store, update and retrieve concepts across both domains.

See https://github.com/Voskod for Python3 source code and basic architecture guide; for Java implementation of CHREST with graphical user interface and more documentation see www.chrest.info.

Results

CHREST was able to learn and apply new concepts in the complex real-world domains of literature and music. It required no ad hoc additions to the fundamental architecture in order to deal with domain specific nuances. The detailed breakdown of the categorisation performance is presented in Table 1. CHREST's categorisation performance was substantially above chance – of the 50 tests across 10 categories (implying 5 correct answers by pure chance), 39 were classed correctly. Within modalities, CHREST correctly categorised 24/30 literature works and 15/20 music pieces. Although the test sample was small (30 pieces of literature, 20 music scores), several patterns have emerged. In the literature task, categorisation of the old Masters – Homer, Chaucer and Shakespeare – produced the highest mean confidence scores (0.31, 0.56 and 0.32 respectively), as well as the highest proportion of true predictions. On the music side, Bach and Mozart had the highest mean confidence scores (0.55 and 0.65 respectively), as well as the highest proportion of true predictions. Of the notable mistakes, Scott was often confused with Shakespeare (possibly due to CHREST seeing Scott's prose, but not poetry

Table 1. CHREST categorisation of literature and music works. Numbers in bold signify its highest confidence score on a given test.

		Homer	Chaucer	Shakespeare	Scott	Dickens	Joyce
Homer	*The Iliad (end chapter)*	**0.43**	0.03	0.14	0.16	0.14	0.10
	The Odyssey (fragment 1)	**0.30**	0.04	0.16	0.20	0.18	0.12
	The Odyssey (fragment 2)	**0.31**	0.07	0.20	0.11	0.18	0.13
	The Odyssey (fragment 3)	**0.29**	0.05	0.17	0.16	0.16	0.17
	The Odyssey (fragment 4)	**0.22**	0.08	0.16	0.14	0.18	0.21
Chaucer	*Canterbury Tales (fragment 1)*	0.13	**0.63**	0.10	0.06	0.04	0.04
	Canterbury Tales (fragment 2)	0.11	**0.57**	0.10	0.05	0.07	0.10
	Canterbury Tales (fragment 3)	0.07	**0.46**	0.14	0.10	0.12	0.10
	Book of Duchesse	0.12	**0.57**	0.10	0.07	0.06	0.07
	The Parliament of Fowles	0.17	**0.56**	0.10	0.07	0.05	0.05
Shakespeare	*Sonnets*	0.20	0.09	**0.37**	0.17	0.06	0.11
	The Passionate Pilgrim	0.31	0.06	**0.30**	0.10	0.15	0.09
	Venus and Adonis	0.21	0.06	**0.32**	0.16	0.12	0.13
	Comedy of Errors	0.14	0.08	**0.31**	0.17	0.22	0.09
	Pericles Prince of Tyre	0.24	0.07	**0.30**	0.11	0.20	0.09
Scott	*Marmion (fragment 1)*	0.19	0.06	**0.23**	0.19	0.18	0.15
	Marmion (fragment 2)	0.17	0.06	**0.28**	0.15	0.19	0.15
	Talisman (fragment 1)	0.15	0.09	0.18	**0.24**	0.20	0.15
	Talisman (fragment 2)	0.18	0.06	**0.24**	0.19	0.20	0.13
	The Black Dwarf	0.15	0.10	0.16	**0.23**	0.19	0.17
Dickens	*A Christmas Carol*	0.20	0.06	0.12	0.22	**0.26**	0.15
	Tale of Two Cities	0.22	0.05	0.19	0.13	**0.28**	0.14
	Oliver Twist (fragment 1)	0.12	0.05	0.18	0.16	**0.30**	0.19
	Oliver Twist (fragment 2)	0.12	0.06	0.19	**0.21**	**0.21**	**0.21**
	The Pickwick papers	0.13	0.08	0.15	0.18	**0.26**	0.19
Joyce	*A Portrait of the Artist (fragment 1)*	0.18	0.02	0.12	0.15	**0.27**	0.26
	A Portrait of the Artist (fragment 2)	0.18	0.03	0.18	0.17	0.20	**0.26**
	Ulysses (end chapter)	0.22	0.01	0.11	0.20	**0.24**	0.22
	Finnegan's Wake (chapter 1)	0.17	0.15	0.17	0.13	**0.19**	**0.19**
	Finnegan's Wake (chapter 2)	0.18	0.07	**0.21**	0.20	0.15	0.19

		Bach	Mozart	Beethoven	Chopin
Bach	*WTC II 21A*	**0.52**	0.13	0.20	0.15
	WTC II 21B	**0.53**	0.11	0.20	0.16
	WTC II 22A	**0.60**	0.11	0.15	0.14
	WTC II 22B	**0.59**	0.11	0.17	0.13
	WTC II 23A	**0.50**	0.10	0.21	0.19
Mozart	*Sonata n10 1mov*	0.27	0.24	**0.31**	0.17
	Sonata n10 2mov	0.09	**0.65**	0.16	0.10
	Sonata n11 full	0.10	**0.69**	0.16	0.04
	Sonata n12 1mov	0.06	**0.87**	0.05	0.02
	Sonata n17 1mov	0.06	**0.81**	0.11	0.02
Beethoven	*Sonata Pathetique 1mov*	0.29	0.07	**0.45**	0.14
	Sonata Pathetique 3mov	**0.39**	0.11	0.35	0.15
	Sonata Moonlight 1mov	0.18	0.04	**0.52**	0.25
	Sonata Moonlight 3mov	**0.38**	0.10	0.28	0.24
	Sonata Waldstein 1mov	0.28	0.11	**0.44**	0.17
Chopin	*Nouvelle Etude n1*	**0.34**	0.08	0.25	**0.34**
	Etude Op.10 n1	0.21	0.08	**0.38**	0.33
	Nocturne Op.9 n1	**0.41**	0.04	0.25	0.30
	Ballad Op.32	0.26	0.16	**0.39**	0.19
	Waltz in A-minor	0.13	0.17	0.30	**0.39**

during training) and Joyce was twice mistaken for Dickens. On one occasion Mozart was confused with Beethoven, Beethoven was twice mistaken for Bach, while Chopin was mistaken for both Bach and Beethoven.

There were no mistakes across modalities – literature was never categorised as music and vice versa. This implies that while the model was taught to classify 10 types of regularities, it has formed (empirically) distinct clusters of chunks that separate the domains of music and literature. Not only was this evident from the overall winning confidence scores, but also from the absence of *any* "LTM chunk activations" across mismatching modalities. To put it another way, there were no cases where a stimulus was, for example, 60% likely to be Mozart, but 0.03% likely to be Homer.

By the end of training, CHREST LTM had developed around 56,000 chunks that encoded both modalities, with the literature modality spanning over 34,000 chunks and music clusters of the LTM taking over 22,000 chunks.

Discussion

There are several key strengths and contributions of the current study. Firstly, computational methodology allowed for rigorous and objective investigation of the fundamental learning mechanisms implicated in human concept formation. Secondly, this model moved away from relying on hand engineered features/dimensions while learning from complex real-life data from multiple modalities. Thirdly, and unlike some pure computer-science machine-learning algorithms, the CHREST architecture is rooted in decades of research in cognitive psychology. The latter point addresses one possible criticism of the current study – despite its lack of comparison with human data, the current findings are still relevant to psychology as they can be viewed as a rigorous extrapolation based on prior research into the cognitive apparatus. Another strength of this study is intuitive realism with regards to its parsimony with training data – only a fraction of Homer's *Iliad* and several Mozart's *Sonatas* were needed to learn generalisable concepts of Homer and Mozart respectively. At this point we should note that research into chess expertise has shown that over 300,000 LTM domain-specific chunks are needed for experts to perform true to their name (Gobet & Simon, 1998; Richman et al., 1996). The LTM volume of the current model was way below that, which may possibly explain some of its performance.

Of the two modelled domains, music does seem to be more impressive. This is in part because music vocabulary and semantics are so abstract and elusive, with music LTM having no intuitively easy moments, as opposed to literature ("doth"? it must be a Shakespeare chunk!). However, this model suffers from an important shortcoming – the dismissed timings in music. While possibly less crucial with regard to the rhythmically rigid classical period composers like Bach and Mozart, romantic period of later Beethoven (and certainly Chopin) relied on tuplet/contrametric rhythm. One extension to the current study would be to model a rhythm LTM network that would run in parallel to other chunking hierarchies.

Another potential theoretical weakness is that author/composer categories were a-priori labelled and predetermined, with learning being supervised and closed to unsupervised clustering of input examples. There are three ways of answering this criticism. Firstly, CHREST architecture has unsupervised learning at its core: the automatic clustering of patterns into chunks is independent of a-priori labels. As mentioned above, CHREST would categorise an occluded "zLiverzool" pattern as "type A", after having been trained on the labelled training data "Liverpool = Type A". However, this very same occluded "zLiverzool" stimulus would also trigger it to recall "Liverpool" if CHREST was trained without any supervision or labels (i.e. if the training set contained just the unlabelled "Liverpool" pattern). Secondly, despite music/literature distinction not being explicitly taught – there were no "literature" or "music" labels during training – CHREST has shown no LTM chunk activation across mismatching modalities. This unsupervised clustering of chunks may point to the creation of distinct concepts of literature and music. Indeed, while CHREST's categorisation confidence scores were a sliding scale that incorporated authors *or* composers, they never incorporated *both*. Thirdly, human readers and musicians do tend to know the name of the author/composer that they are studying, thus largely matching the labelled stimuli approach of this concept formation study.

A common general criticism of the computational modelling approach is the potential for "overfitting" – changing free parameters to achieve better fit may lead to poor generalisability beyond the currently simulated data (Tetko, Livingstone, & Luik, 1995). This study followed Simon's (1992) advice and attempted to address the issue by doubling the *data explained/free parameters used* ratio – the same free parameters were used for both literature and music. However, conclusive guarantee that the model provides a unique explanation is impossible (Lakatos, 1970).

With these qualifiers out of the way, Chunking Theory does seem to provide a general insight into the psychology of concept formation beyond the two modelled domains. Utilising the rigour of a formal model, the CHREST architecture connects fundamental psychological structures such as LTM/STM to the detailed ground up process of learning to categorise – a framework that can potentially be applied to *any* symbolic domain. Furthermore, the current study also shows CHREST's power to operationalise the factors underpinning subjectivity – learning a concept is a function of many potentially unique variables. The variation in prior knowledge, the amount of data and learning cycles devoted to learning a concept, the order in which this data is learnt and the agent's internal parameters (including the span of the attention window, STM size, the likelihood and speed of forming or updating a chunk) – are all part of CHREST's computational methodology and all play a part in predicting how individuals come to share subjective states… Like concepts of Mozart or Homer.

We may thus conclude by paraphrasing Herbert Simon (1992): it is a justified conclusion that human concepts can

be characterised by operations in STM/LTM with CHREST-like architecture, although the detailed structure of the model is open to further enrichment.

References

Barsalou, L. W. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes, 18*(5-6), 513.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1521), 1281.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55-81.

De Groot, A. D. (1965). *Thought and choice in chess.* The Hague: Mouton Publishers.

Gobet, F., & Lane, P. (2012). Chunking mechanisms and learning. In M. M. Seel (Ed.), *Encyclopedia of the sciences of learning*. New York: NY: Springer.

Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences, 5*(6), 236.

Gobet, F., & Simon, H. (1996c). Templates in chess memory: a mechanism for recalling several boards. *Cognitive Psychology, 31*, 1-40.

Gobet, F., & Simon, H. (1998). Expert chess memory: revisiting the chunking hypothesis. *Memory, 6*, 225-255.

Goldstone, R. L., & Steyvers, M. (2001). The sensitization and differentiation of dimensions during category learning. *Journal of Experimental Psychology: General, 130*(1).

Hampton, J. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior, 18*, 441-446.

Hull, C. L. (1920). Quantitative aspects of evolution of concepts. *Psychological Monographs, 28*.

Lakatos, I. (1970). Falsification and methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge*. Cambridge: CUP.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Murphy, G. L. (2002). *The big book of concepts*: Cambridge, MA: MIT Press, 2002.

Nosofsky, R. M. (2011). The generalized context model: An exemplar model of classification. In *Formal Approaches in Categorization.* (pp. 18-39). New York, NY, US: Cambridge University Press.

Nosofsky, R. M., Sanders, C. A., & McDaniel, M. A. (2018). Tests of an exemplar-memory model of classification learning in a high-dimensional natural-science category domain. *Journal of Experimental Psychology: General, 147*(3), 328-353. doi:10.1037/xge0000369

Richman, H. B., Gobet, F., Staszewski, J., & Simon, H. (1996). Perceptual and memory processes in the acquisition of expert performance: The EPAM model. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 167–187). Mahwah, MA: Erlbaum.

Richman, H. B., Simon, H., & Feigenbaum, E. A. (2002). Simulations of paired associate learning using EPAM VI. *Unpublished*.

Richman, H. B., Staszewski, J. J., & Simon, H. A. (1995). Simulation of expert memory using EPAM IV. *Psychological Review, 102*(2), 305-330.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*, 192-233.

Simon, H. A. (1974). How big is a chunk? *Science, 183*(4124), 482-488.

Simon, H. A. (1992). What is an "explanation" of behavior? *Psychological Science, 3*, 150-161.

Smith, D. J., & Minda, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(1), 3-27.

Tetko, I. V., Livingstone, D. J., & Luik, A. I. (1995). Neural network studies. 1. Comparison of overfitting and overtraining. *Journal of Chemical Information and Computer Sciences, 35*(5), 826.

Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Blackwell.