Using a Cognitive Architecture for Addressing the Question of Cognitive Universals in Cross-Cultural Psychology: The Example of Awalé

Fernand Gobet

Journal of Cross-Cultural Psychology 2009; 40; 627 originally published online May 14, 2009; DOI: 10.1177/0022022109335186

The online version of this article can be found at:
http://jcc.sagepub.com/cgi/content/abstract/40/4/627

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
International Association for Cross-Cultural Psychology

Additional services and information for Journal of Cross-Cultural Psychology can be found at:

Email Alerts: http://jcc.sagepub.com/cgi/alerts

Subscriptions: http://jcc.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations http://jcc.sagepub.com/cgi/content/refs/40/4/627
Using a Cognitive Architecture for Addressing the Question of Cognitive Universals in Cross-Cultural Psychology

The Example of Awalé

Fernand Gobet
Brunel University

A central theme in cross-cultural psychology is the extent to which cognitive mechanisms are universal, or, alternatively, are specific to a given culture. A new way to tackle this question is proposed: to use the same cognitive architecture, implemented as a computer program, for simulating phenomena in which individuals from different cultures perform a task familiar to their own culture. The CHREST architecture has simulated a number of empirical phenomena related to the Western board game of chess. Here, we show that a model implemented in the same architecture accounts for several phenomena in awalé, a board game from the mancala family, which is commonly played in western Africa and in the Caribbean. CHREST first learns chunks by scanning expert-level games and then is placed in memory experiments and problem-solving situations similar to those used with human youngsters. The model replicates empirical phenomena on memory for awalé positions reasonably well, although not perfectly, and also learns to play a fair, but far from perfect game using pattern recognition. The assumptions that learning is mediated by the acquisition of a large number of chunks and that the capacity of visual short-term memory is limited to three chunks are important in explaining the empirical data for the two games. The implications for theory development in cross-cultural psychology are discussed.

Keywords: awalé; chess; chunking; cognitive architecture; cross-cultural psychology; expertise; short-term memory; universality of processes

A key question in cross-cultural psychology is the extent to which psychological processes are universal or specific to particular cultures (Kagitcibasi & Berry, 1989; Lehman, Chiu, & Schaller, 2004; Yamagata et al., 2006), where culture can be defined as “the shared

Author’s Note: Part of this research was presented at the colloquium “Mancala Games: New Perspectives,” held in 2006 at the University of Fribourg (Switzerland). We thank Jean Retschitzki for inviting us to this colloquium and providing us with the database of awalé games, as well as Richard Ll. Smith and Yvan Russell for comments on drafts of the manuscript. Please address correspondence to Fernand Gobet, Centre for the Study of Expertise, Brunel University, Uxbridge, Middlesex, UB8 3PH, United Kingdom; e-mail: Fernand.Gobet@brunel.ac.uk.
way of life of a group of people” (Berry, Poortinga, Segall, & Dasen, 1992, p. 1). This question has been studied in a number of domains, including face perception (Izard, 1994), development (Haight, Wang, Fung, Williams, & Mintz, 1999), and personality (Yamagata et al., 2006). An influential way to address this issue has been to study individuals performing tasks specific to their own culture and to investigate whether the processes engaged, for example cognitive processes, are the same in different cultures. In particular, play and games have often been used to study the universality of cognitive and developmental mechanisms, as these activities can be found in many non-Western cultures, as opposed to activities such as those carried out in a schooling environment.

Perhaps with the exception of Piaget’s theory, which has offered a unified framework for a fair number of cross-cultural studies (e.g., Dasen, Inhelder, Lavallée, & Retschitzki, 1978; Triandis & Heron, 1981), cross-cultural researchers have tended to eschew highly formal theories for addressing cognition across cultures, perhaps given the difficulty to apply them to activities that are shaped by different cultural contexts and thus may appear rather dissimilar—at least superficially. In particular, we could find no study where computer modeling was used to study the question of the universality of processes. This is a regrettable methodological omission, as theories expressed as computer programs offer more flexibility than for example mathematical or logical formalisms (Newell & Simon, 1972). In addition, some recent computational theories, sometimes known as unified theories of cognition or cognitive architectures, have been able to simulate diverse aspects of behavior and provide the necessary generality to tackle the issues linked to the question of the universality of cognitive processes. Examples of such theories are ACT-R (Adaptive Control of Thought–Rational; Anderson et al., 2004), Soar (Newell, 1990), and CHREST (Chunk Hierarchy and REtrieval STructures; Gobet, 2001; Gobet et al., 2001). The purpose of this article is then to show that a specific cognitive architecture, CHREST, can shed important light on whether the same basic cognitive mechanisms can explain a number of phenomena in two board games, one Western game (chess) and one game of African origin (awalé). As will be detailed below, CHREST has already accounted for numerous data on chess expertise (e.g., Gobet, 1998; Gobet, de Voogt, & Retschitzki, 2004). Therefore, the focus of the article will be on awalé and the goal will be to show that the same mechanisms underlying chess expertise according to CHREST also underlie awalé expertise.

Beyond their interest for cross-cultural psychology, board games have provided invaluable information in the study of human cognition. Some of the important concepts of cognitive science, such as progressive deepening, selective search, and pattern-based recognition, originate from this research, and in particular, from research into chess (Gobet et al., 2004). A substantial number of empirical results have also been collected in this field, and this body of empirical results has enabled the development of precise theories implemented as computer programs. In particular, the tradition of information-processing models started by Simon and Barenfeld (1969) has led to the CHREST architecture, which has simulated a large amount of data on chess expertise (see section “Chunking Mechanisms and the CHREST Cognitive Architecture,” below). The basic assumption of CHREST is that learning resides in the acquisition of units of perception and meaning, known as chunks. Beyond chess, CHREST has explained phenomena in two other domains of expertise: computer programming (Gobet & Oliver, 2002) and physics (Lane, Cheng, & Gobet, 2000). Beyond expertise,
the architecture has been used to successfully simulate aspects of concept formation (Gobet, Richman, Staszewski, & Simon, 1997; Lane & Gobet, 2005), children’s acquisition of syntactic categories in four different languages (Freudenthal, Pine, & Gobet, 2006; Freudenthal, Pine, Aguado-Orea, & Gobet, 2007), and children’s acquisition of vocabulary (Jones, Gobet, & Pine, 2007, in press).

In spite of a large empirical support for the CHREST architecture in the domain of chess, a western board game in its modern form, it is unknown whether the same mechanisms and parameters could apply to board games that are played in other cultures and that follow totally different rules. In this respect, mancala games offer an ideal way to test the generality of the mechanisms postulated by CHREST: the games of this family are typically played in Africa, Southeast Asia, and the Caribbean, where they fulfill an important societal and cultural role. Furthermore, the psychology of some mancala games has been studied scientifically, most notably the game of awalé played in West Africa and the Caribbean (Retschitzki, 1990; Retschitzki, Keller, & Loeschberger, 1984; Retschitzki, Loeschberger, Gut, & Brulhart, 1986) and bao, played in Zanzibar (De Voogt, 1995, 2002).

It is not obvious a priori that the mechanisms and parameters so successful in explaining chess players’ behavior—in particular the mechanisms related to chunking and the assumption that visual short-term memory (STM) has a limited capacity of three items—also apply to mancala games. Differences in the cultural environment may lead to differences in the way certain cognitive mechanisms are used for specific tasks, and it would not be implausible that different cultures lead to the use of different learning strategies. Importantly, it has been proposed that the cognitive processes of African mancala players are not comparable with those of (occidental) chess players (De Voogt, 1995, 2002). Given the vast differences in the cultural contexts between chess and awalé, the mancala game most studied in the scientific psychology literature, applying CHREST to the latter can be seen as a strong test of the theory. The article first describes the game of awalé and presents the main psychological findings that have been obtained on this game. It then reviews the main theoretical assumptions behind CHREST, and presents some of the key empirical support from chess. This sets up the scene for the presentation of how the architecture has been used to develop an awalé model. Simulations of experiments on memory and problem solving are then presented. The discussion evaluates the extent to which the model accounts for key results successfully and discusses the implications for the study of cognitive universals in cognitive psychology.

The Game of Awalé

Awalé derives its name from the Baoulé name for the plant Calsapinia crista, as it is usually played with its seeds. It is typically played by men, outdoors and in public. Players are usually surrounded by a crowd of onlookers, who do not hesitate to comment loudly not only on the game but also on current affairs of interest to them. Thus, awalé plays an important societal role in the African countries where it is played. Retschitzki (1990) notes that this game has several advantages from the point of view of cross-cultural research. First, it has not been influenced by Westerners and thus is a direct reflection of the local culture. This makes it possible to avoid the methodological problems that affect research using tasks foreign to a specific culture. Second, while using a rich vocabulary of concepts
Awalé belongs to the family of mancala games, a family of board games also called sowing games. These games consist of a set of monochrome seeds and a board containing holes. The game proceeds by picking up the seeds from a hole and then sowing them one at a time in the following holes. Captures are made as a function of the state of board after the end of sowing (or in some games during sowing). The goal is to capture more seeds than the opponent. The game has different names depending on the country where it is played. These include wari, oware, ayo, ouri, and awari. We will use the name “awalé,” the appellation used in Ivory Coast, as the experiments we will describe later have been carried out in that country.

There are a number of variations of awalé, but this description focuses on the variation used later in this article. The board consists of two rows of six holes and there are 48 pieces, commonly called seeds. (See Figure 1, upper part, for the position of the pieces at the beginning of the game, and Figure 1, lower part, for a middle-game position.) Each player owns one row. As in other mancala games, the purpose of the game is to capture as many seeds as possible.

A move consists in selecting the contents of one hole on the side of the player and distributing the seeds one by one in subsequent holes counterclockwise, until all seeds have been spread. Captures occur when two conditions are met: the last seed enters a hole on the opponent’s row, and there are a total of two or three seeds in the last hole. When the two
conditions are met, the seeds in the last hole are captured and taken from the board. If the holes on the opponent’s row directly preceding the hole where the capture occurred also contain two or three seeds, these are also captured. Seeds are not spread in the hole from which the sowing originated (this rule applies when a hole contains more than 11 seeds). If a player has no seed, the opponent must play a move that gives her at least one seed. If this is not possible, the game ends, and the opponent adds the remaining seeds (minus one seed that is given to the other player) to the captured seeds. If the only move a player can play would capture all the remaining seeds of the other player, who then could not play, this move is carried out but no seeds are captured. Readers interested in the differences in rules between variants of awalé are referred to Bell (1960), Murray (1952), and Retschitzki (1990).

The Cognitive Psychology of Awalé Playing

Retschitzki and colleagues (Retschitzki, 1990; Retschitzki et al., 1984; Retschitzki et al., 1986) carried out a systematic study in Ivory Coast on the determinants of cognitive ability in awalé. The study took place in the village of Kpouébo, which is inhabited by members of the Baoulé ethnic group. In the main component of the study, Retschitzki submitted 38 male Baoulé youngsters (from 9 to 15 years old) to a variety of tests, either related or unrelated to awalé, and measuring memory, estimation of quantity, and arithmetic knowledge. An additional variable under study was whether the material was real (e.g., real board and seeds for awalé) or symbolic (e.g., use of numbers with awalé). Retschitzki also provided a detailed analysis of the evolution of strategies and types of reasoning shown by youngster and adult awalé players. (The results of the memory experiments will be presented in the section “Memory for Awalé Positions,” below.)

N’Guessan (1992) carried out a study of the acquisition of methods of play (tactics, strategies) in awalé and evaluated different methods of teaching awalé to Swiss participants. Like Retschitzki (1990), and in line with previous work on chess (Simon & Chase, 1973), N’Guessan reached the conclusion that skill in awalé requires substantial practice, as it involves the acquisition of a large amount of well-organized knowledge and is a combination of perceptual abilities and mastery of numerous strategies. As the development of skill appeared to follow the same path with Swiss players as with Baoulé players, N’Guessan also concluded that his study supported the hypothesis of the universality of learning mechanisms.

Several computer programs playing awalé have been developed, based on standard artificial-intelligence search algorithms, as discussed by Retschitzki (1990) and N’Guessan (1992). A variant of the game has been solved computationally, with perfect play leading to a draw (Romein & Bal, 2002). There also exist some programs aimed at simulating aspects of human cognition. Retschitzki (1990) presents simulation models choosing a move based on a choice of offensive or defensive play and on the type of threats noticed. However, there exists no model in the literature simulating awalé players’ memory and providing an explanation for the link between high-level perception and decision making in that game. As noted above, the aim of this article is to fill in this gap by using a general cognitive model based on chunking.
The CHREST Cognitive Architecture

Inspired by Simon and Chase’s (1973) chunking theory, the CHREST architecture (Gobet & Simon, 1996b, 2000) has closely simulated several phenomena in chess expertise and in other domains (for reviews, see Gobet & Lane, 2005; Gobet et al., 2001). The architecture combines low-level aspects of cognition (e.g., mechanisms monitoring information in STM) with high-level aspects of cognition (e.g., use of strategies). Learning is seen as the acquisition of a network of long-term memory (LTM) nodes, also called chunks, which can become connected as a function of the similarity of their contents. Patterns that recur often in the environment make it possible for chunks to evolve into more complex data structures, known as templates. Templates are schema-like structures that consist of a core pattern with peripheral slots which can vary without altering the core (Gobet & Simon, 1996b). The core and slot structure allows newly encountered patterns to be encoded and categorized rapidly. The architecture also makes assumptions about the capacity of visual STM (limited to three chunks). Time parameters are taken into account for each process of the architecture (e.g., it takes 8 seconds to create a new chunk, Simon & Chase, 1973). Simulations are carried out by letting the computer program acquire knowledge by receiving stimuli representative of the domain under study. For example, during the learning phase of the chess simulations, the program incrementally acquires chunks and templates by scanning a large database of positions taken from master-level games. This makes it possible to create nets of various sizes and so to simulate the behavior of players of different skill levels. Taken together with the presence of time and capacity parameters, this enables CHREST to make unambiguous and quantitative predictions. A considerable number of data have been accounted for by CHREST with respect to chess expertise, from novice to grandmaster (De Groot & Gobet, 1996; Gobet, 1997; Gobet & Jackson, 2002; Gobet & Simon, 1996a, 2000; Gobet & Waters, 2003; Waters & Gobet, 2008). With perception, CHREST has simulated the eye movements during the first 5 seconds of the presentation of a position as well as the rapid recognition of chunks and templates. With memory, CHREST accounts for the effect of various types of position modification and randomization, the role of presentation time in memory, the type of errors made, and the type of chunks replaced. It can also simulate the detail of how novices acquire chunks and templates. With problem solving, the theory explains the ability of strong players to identify potential moves rapidly and correctly and also the fact that the average depth of search increases as a function of skill with diminishing returns, following a power function. A recent review of the literature has shown that the CHREST model offers the best theoretical explanation for the large body of experimental and observational data on chess (Gobet et al., 2004).

Description of the Awalé Model

Based on CHREST’s pedigree, and in particular its track record in simulating chess expertise, the architecture seems to be in a good position to simulate data on awalé. At the same time, as noted in the introduction, awalé differs from chess in important ways, with De Voogt (2002) arguing that different cognitive mechanisms underpin the way Africans...
play mancala games. A successful application of CHREST to awalé would thus have theoretical implications well beyond board game studies.

Our general approach will be the same as that used with chess (Gobet & Jansen, 1994; Gobet & Simon, 2000; Gobet & Waters, 2003). First, we let the program learn from a set of expert-level games and, second, we put the model in the same situation as human players, both for a memory and a decision making task. In the former case, we will be particularly interested whether the capacity of visual STM postulated by CHREST—three chunks—holds in this new domain as well. In the latter case, we will focus on the situation where the problems have to be solved rapidly by pattern recognition, that is, without enabling the model to use search. As awalé games are played extremely rapidly (Retschitzki, 1990)—much more so than chess games—the focus on rapid play is appropriate.

As with chess, the model uses its eye movements to determine what is learnt. We assume that four anchor points are fixated, as shown in Figure 2. The four anchor points are Front-a (Fa), Front-b (Fb), Back-a (Ba), and Back-b (Bb), and they follow the direction of play (counterclockwise). At the beginning, chunks are learnt from the anchor points. For example, referring to Figure 2, the model would learn (Fa 1) first, then (Fa 1 1), then (Fa 1 1 6), or (Fb 0), then (Fb 0 0), then (Fb 0 0 1). After a while, chunks can be combined to form larger chunks. For example, the two chunks we have just described could be combined in the super-chunk ((Fa 1 1 6) [Fb 0 0 1]).

When attempting to recognize a chunk, CHREST outputs the largest it can, given an anchor point. So, assuming that CHREST knows the super-chunk ((Fa 1 1 6) [Fb 0 0 1]), fixating the first anchor point (Fa), and then the second (Fb), outputs the super-chunk ((Fa 1 1 6) [Fb 0 0 1]), and not the subelements of this super-chunk, such as (Fa 1 1 6) or (Fa 1 1).

A key assumption of the model is that a move can be associated to a chunk, forming what is known as a production (Newell & Simon, 1972). So, referring to Figure 2, if South decides to sow the seeds from the third hole from the left (hole #3), the program could learn, depending on its stage of learning, one of the following productions:

1. (Fa 1 1) → select the seeds in hole #3
2. (Fa 1 1 6) → select the seeds in hole #3
3. ((Ba 0 0 1) (Bb 0 1 0)) → select the seeds in hole #3

If (b) is learnt, the next time CHREST recognizes the chunk (Fa 1 1 6), it would intuitively suggest choosing hole #3 as a possible move. With human players, further look-ahead...
search may sometimes be used to confirm or disconfirm this choice, depending on the situation of the entire board.

Material Used for Simulating Learning

Unlike chess and other western games (such as checkers), where players routinely write down their moves when playing a game, awalé players follow the custom of other Mancala players of not recording their moves. This is likely to be due both to the speed with which games are played and the lack of an agreed-on notational system (Gobet et al., 2004). Thus, contrary to games such as chess for which there exist large databases of games played by experts, the number of awalé games that have been transcribed is very small indeed. We could find only 80 games that were played in the international tournaments of Antigua in 2004 and La Tour-de-Peilz (Switzerland) in 2005, making a total of 8,059 different positions. (In contrast, chess simulations with CHREST typically use databases of 50,000 positions.) In these tournaments, the games were videotaped as part of a research project on the psychology of awalé carried out by Jean Retschitzki at the University of Fribourg (Switzerland) and then transcribed.

Another issue, also related to the cultural environment in which awalé is played, is that there is no official system of titles. Rather, players use informal definitions of expertise, based on players’ reputation or their results in some specific tournament (see Gobet et al., 2004, for details). Similarly, and again unlike chess, there is no rating associated with each player that allows an unambiguous measure of players’ strength. However, taking into account the results of several tournaments, there is no doubt that the players whose games were used to train CHREST were among the best in the world.

During learning, CHREST scanned each position of the training set using the anchor points and learned the respective chunks and associated moves. The program also kept track of how often a chunk had been associated with a given move. During testing, it was placed in the same situation as a human, as will be described below. Our focus was in testing the generality of the chess model by applying its key assumptions with minimal change to simulating awalé phenomena. Therefore, we have not fitted any parameters to increase the quality of fit between model and data. Additional free parameters would undoubtedly improve the goodness of fit in some simulations, but at the cost of theoretical transparency.

Memory for Awalé Positions

Human Data

The first set of simulations focuses on the memory experiment carried out by Retschitzki and colleagues (Retschitzki, 1990; Retschitzki et al., 1984). Youngsters from 9 to 15 years of age participated in this experiment and a test measuring knowledge of awalé rules and simple tactics was used to allocate the participants to two skill levels: average and strong.

Retschitzki used two conditions, symbolic and real. With symbolic material, a position was presented for 30 seconds, with three types of representation: black dots representing the seeds with circles representing the holes; numbers in circles representing the number of
seeds in a hole; and photographs of actual boards. There were five types of positions: unitary (there is at most one seed in any hole), neutral (no player threatens a capture), simple threat (one of the two players threatens to win material in one move), reciprocal threat (both players threaten to win material in one move), and double threat (one player threatens to win material in two different ways). In the real material condition, a real board was used and the position to memorize could either come from a game being played by the participant or a position out of context. The presentation with real positions was 15 seconds.

Retschitzki found several important results with the symbolic material, all three types of representation being pooled. First, performance increased as a function of age. Second, performance increased as a function of skill. Third, positions with more squares occupied tended to be harder to memorize. And fourth, there were no differences as a function of the type of position. With the real positions, there was no effect of age or skill level for the positions taken from a game being played, whereas there was a skill effect with the positions taken out of context.

Methods

Scoring of recall. The simulations will focus on the conditions with symbolic material, and, as Retschitzki (1990) averaged the results of the three formats used (photographs, diagrams with dots denoting seeds, and diagrams with numbers representing the number of seeds), we will not differentiate between them. The scoring system was the same as Retschitzki’s. The correspondence between each recalled position and its stimulus was evaluated using three criteria: 1 point was allocated for each occupied hole that was in fact occupied in the stimulus (irrespectively of the number of seeds), 1 point was allocated for each correct number of seeds (irrespectively of the hole where they were placed), and 1 point was allocated for choosing the correct hole together with the correct number of seeds. A penalty was also given when the number of occupied holes was overestimated, with -1 point for each hole that was deemed occupied while in fact it was not. To make this measure comparable between positions having a different number of seeds, the total obtained with the criteria just described was divided by the maximum number of points possible for a given position (i.e., the number of occupied holes multiplied by 3).

Training and testing the model. The model was trained using 5,000 positions randomly chosen from the set of 8,059 different positions mentioned above. To get an estimated of variability, this procedure was carried out 500 times. The test positions were the same as those used by Retschitzki (1990). During the test phase, the model fixated the four anchor points and attempted to recognize a chunk for each of them. These chunks were put into STM until its capacity was filled. To keep the simulations simple, learning and the use of templates were turned off during the testing phase. (In the simulations described in the next section, where the aim was to find the most suitable capacity for STM, the first four STM slots were filled using the chunks obtained by fixating the anchor points, and additional slots were filled by adding or correcting the values held in the first four slots.)

Setting the development parameter and the short-term memory capacity parameter. So that simulations can be carried out, four parameters must be set as follows: (a) a development
parameter, making it possible to have different simulated ages; (b) STM capacity; (c) number of LTM chunks for the average group; and (d) number of LTM chunks for the strong group. A possible approach would have been to perform an exhaustive search in the space of parameters and choose the parameters best fitting the data. Because we were interested in testing the generalizability of CHREST, we preferred a second, more theoretical approach. For the capacity of visual STM, we used the parameter already present in CHREST (three chunks), but also verified that this estimate was reasonable (see below). As for the development parameter, we estimated the parameter based on independent data, as discussed in the next paragraph.

Retschitzki (1990) carried out a series of control tasks, including a quantity estimation task in which the participants had either to estimate the number of dots (from 3 to 15) on a card presented for 3 seconds, or the number of seeds (from 4 to 15) in a hole the content of which was visible for 3 seconds as well. Estimation of the number of dots increased linearly as a function of age, whereas estimation of seeds had an inverted-U relationship with age (see Retschitzki, 1990, figures 4.11 and 4.13). The data on estimation of quantity were used to compute the probability of recognizing a chunk by dividing the score at a given age by the score at age 15 years and combining the results between the two tasks (see Table 1). For example, the likelihood of recognizing a chunk at age 9 years is .71, as compared with 1.00 at age 15 years. In the simulations, these probabilities were used to modulate recall, the rationale being that incorrectly evaluating the number of seeds in a hole impairs chunk recognition and thus recall.

For the number of LTM chunks characterizing the average and strong groups, we decided that they should be independent from age (i.e., the average players will have the same number of chunks in the four age levels and the same will apply to the strong players). The value assignment was based on the performance of the 15-year-old players. Figure 5 (top) shows the proportion of correct recall as a function of age and skill level. This information is also provided in the four panels of Figure 3, where memory capacity was systematically varied from 1 to 9, for the four probabilities shown in Table 1. As they are based on 500 independent replications, these simulations are highly reliable. For a capacity of 3 and a probability of 1.00, corresponding to the 15-year-old players, the best fit is obtained with 100 positions for the average players (equivalent to 175 chunks) and 5000 positions for the strong players (equivalent to 7855 chunks). Thus, skill level was manipulated by using one version of the model with 175 chunks and another version with 7855 chunks.

Two points may be noted at this juncture. First, the human data are fairly noisy: the 11-year-olds do not show any skill differences, and the 13-year-olds perform better than the 15-year-olds. Retschitzki (1990) notes these two anomalies, but does not provide any explanation for them. Second, Figure 3 shows that the capacity of 3 chunks, which was

### Table 1

Estimated Probability of Recognizing a Chunk as a Function of Age

<table>
<thead>
<tr>
<th>Age (Years)</th>
<th>Estimated Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>0.71</td>
</tr>
<tr>
<td>19</td>
<td>0.93</td>
</tr>
<tr>
<td>13</td>
<td>0.96</td>
</tr>
<tr>
<td>15</td>
<td>1.00</td>
</tr>
</tbody>
</table>
selected for theoretical reasons, turned out to be a good choice. Compare this with the values 1 and 2 for STM capacity. These yield curves that never reach the levels of the average players, whereas the high values from 5 to 9 move too rapidly, with few chunks, up to the level of average players, and then rapidly onto the level of strong players. This pattern is not plausible for a game as complex as awalé. The only realistic alternative is an STM of 4 (with a very similar pattern to 3), but for theoretical reasons (Gobet & Simon, 2000; Waters & Gobet, 2008) we preferred the capacity of 3.

Results

General pattern of learning. The top panel of Figure 4 shows how memory performance increases as a function of the number of positions seen (the development parameter was set
to 1.00) and the bottom panel shows how memory performance increases as a function of the number of chunks acquired. Both graphs show the same pattern, with recall performance being well-fitted by a logarithmic function ($r^2 = .70$, $p < .01$, for the number of positions, and $r^2 = .73$, $p < .01$, for the number of chunks). There is indeed a strong linear correlation between the number of positions seen and the number of chunks learnt ($r = 1.00$, $p < .001$). The same pattern has been found with chess simulations (Gobet, 1993b). Thus, whereas the learning of new chunks increases linearly with the number of positions seen, performance increases with diminishing returns. The explanation for this counterintuitive pattern of results is the same as that proposed by Newell and Rosenbloom (1981) for Soar: with experience in a domain, increasingly larger chunks are learnt, but because larger chunks impose more constraints than smaller ones, the likelihood of using them becomes increasingly smaller.

Note: The standard errors of the mean were all inferior to .01 and are not shown. The simulations are based on 500 replications.
Effects of age and experience. As shown in Figure 5, the simulations yielded a main effect of age as well as a main effect of experience, just like in the human data. The goodness of fit is reasonably good \( r^2 = .81, p < .005; \) root mean square error [RMSE] = 0.073). This result must be qualified by the fact that the data for the 15-year-olds were fitted during the search for reasonable values for the number of LTM chunks for average and good players.

*Number of occupied squares.* Figure 6 shows memory performance as a function of the number of holes that contain seeds. To be consistent with Retschitzki’s (1990) presentation, we have combined the results of the average and strong players. With the human data, performance gets slightly worse as the number of occupied holes increases, although a blip
occurs with six holes. The simulations provide a good fit for the cases with 5, 6, and 7 holes, both with respect to the pattern of results and the absolute values ($r^2 = .82$, RMSE = 0.028), but totally underestimate recall with 4 holes, with the consequence that the overall fit is very poor ($r^2 = .01$, RMSE = 0.060).

*Types of position.* Following Retschitzki (1990), we did not include the unitary positions in this analysis. In Retschitzki’s data, no reliable differences were found with the four types of position (see Figure 7, top). Although the fit between simulation and data is reasonable ($r^2 = .50$, $p < .05$, and RMSE = 0.071), this seems mostly explained by the skill differences.

**Figure 6**  
Proportion of Recall as a Function of the Number of Occupied Holes (the Two Skill Levels Are Pooled): (Top Panel) Human Data, (Bottom Panel) Simulations

Note: In the simulations, the standard errors of the mean were all inferior to .01 and are not shown. The simulations are based on 500 replications.
Two outcomes differ between simulations and data. First, the simulated strong players show a small but clear trend to improve their performance across the four types of positions ($r^2 = .99, p < .005$), a pattern of results which is not present with the human data. Second, the simulations with the average players show a decrease in performance with the double-threat positions, a decrease that is not present with the human data.

Rapid Decision Making in Awalé

An interesting feature of awalé is that even competitive games are played very rapidly, and an entire game rarely lasts more than 10 minutes. It is thus plausible that there is not enough time for extensive look-ahead search and that a fair amount of the moves are chosen...
by pattern recognition. Informal discussion with world-class players supported this hypothesis (September Christian and Trevor Simon, personal communication, September 5, 2006). In spite of the relatively limited number of positions used for training, in the simulation below we establish the extent to which the model can select good moves.

Methods

Learning was carried out in the same way as with the memory simulations, with the only difference being that, rather than choosing several networks with an incremental number of chunks to simulate different levels of skill, we selected only one network with full learning of the training positions available.

The experiment simply consisted in showing a set of positions to CHREST and letting it choose a move by pattern recognition (without anticipation). When several moves were proposed for the same position, three different criteria were used for selecting the best move: (a) the number of chunks voting for a move, (b) the sum over all chunks of the frequency of association between chunk and a given move, and (c) the size of the largest chunk proposing a move. With all criteria, the move that ranked first was chosen. We will present results with each of these criteria. When one or more moves were tied, we used two selection criteria and we will present results for both criteria: (a) “Tie” criterion: the target move was one of the moves ranked first and (b) “Only one” criterion: in case of tie, a move was picked up randomly. Note that, given the structure of the game, the chance of finding the best move by chance is 1/6 (.167) when all six holes in the row owned by the player are filled and is higher when the row includes empty holes.4

Simulations

We carried out three variations of the choice-of-move task: (a) finding the best move in positions that were presented during learning, (b) finding the best move in positions taken from the main database that were not presented during learning, and (c) finding the best move in the test set of elementary tactics used by Retschitzki (1990) for assessing players’ strength. In all cases, the simulations were independently replicated 500 times. The model was trained with 7,559 positions randomly chosen from the pool of positions available (CHREST scanned each position only once), yielding networks with 11,249 nodes on average. In all test positions, it was South to move.

Finding the best move in seen positions. In this simulation, 250 positions that were used during learning were randomly selected for the test phase. The number of chunks is the criterion obtaining the best performance, just ahead of the frequency of association (see Table 2). Note that, even with the stricter criterion of randomly picking up a move in case of tie, the program obtains a high performance (.874 with the number of chunks criterion). But why is it that the model does not perfectly remember the best move in the positions it has already seen? The answer is that, just like humans, the model forgets. The structure of the network changes as more chunks are learned, which may lead to the situation where the production containing the best move is by-passed, as there is a larger chunk below it (the model always recognizes the largest chunk possible). Also, a specific position may have several
chunks voting against the best move, and although the best move is one of the candidates, it may be out-voted by the other moves.

Finding the best move in new positions. In this simulation, 250 positions not encountered during learning were selected for the test phase. Not surprisingly, the performance is worse than in the previous simulation (see Table 3). The number of chunks criterion obtained the best result with the tie criterion, whereas the frequency of association criterion obtained the best performance with the only-one criterion.

Finding the best move in simple awalé problems. Eight problems of elementary tactics were taken out of the nine used by Retschitzki (1990). These problems were used to identify the players’ strength in the memory study reported above.

The number-of-chunks criterion obtains the best performance, with both tie criteria (see Table 4). With the stricter only-one criterion, performance is clearly above chance level with the number-of-chunk and largest-chunk criteria, but not so with the frequency-of-association criterion, which does poorly even with the tied criterion. Overall, CHREST’s performance is clearly worse than the youngsters studied by Retschitzki (1990). For example, the 9-year-old

---

Table 2
Proportion of Played Moves Found by CHREST When Attempting to Find Moves in Seen Positions

<table>
<thead>
<tr>
<th>Criterion for Move Selection</th>
<th>Number of Chunks</th>
<th>Frequency of Association</th>
<th>Largest Chunk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tied</td>
<td>.964 (.001)</td>
<td>.920 (.001)</td>
<td>.718 (.001)</td>
</tr>
<tr>
<td>Only one</td>
<td>.874 (.001)</td>
<td>.860 (.001)</td>
<td>.708 (.001)</td>
</tr>
</tbody>
</table>

Note: The results are based on 500 replications. Learning was carried out using 7,559 positions, yielding networks with 11,250 nodes on average (SEM = 1.3). Testing was done on 250 positions. The average probability of finding the correct move by chance is .284 for this set of test positions. Values in parentheses indicate standard errors of the mean (SEM).

Table 3
Proportion of Played Moves Found by CHREST When Attempting to Find Moves in Unseen Positions

<table>
<thead>
<tr>
<th>Criterion for Move Selection</th>
<th>Number of Chunks</th>
<th>Frequency of Association</th>
<th>Largest Chunk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tied</td>
<td>.453 (.001)</td>
<td>.432 (.001)</td>
<td>.275 (.001)</td>
</tr>
<tr>
<td>Only one</td>
<td>.287 (.001)</td>
<td>.323 (.001)</td>
<td>.272 (.001)</td>
</tr>
</tbody>
</table>

Note: The results are based on 500 replications. Learning was carried out using 7,559 positions, yielding networks with 11,248 nodes on average (SEM = 1.1). Testing was done on 250 positions. The average probability of finding the correct move by chance is .284 for this set of test positions. Values in parentheses indicate standard errors of the mean (SEM).
players solved 64% of the problems correctly and the 15-year-old players solved all problems correctly.

**Discussion**

The aim of this article was to provide a new methodology for addressing the question of cognitive universals in cross-cultural psychology. More specifically, the article tested the generality of the CHREST architecture—with a focus on its chunking mechanisms and the capacity of visual STM it postulates—by exploring to what extent it could simulate results on awalé, a non-Western board game that is often played in African and Caribbean countries. To ensure a sufficiently large coverage of the simulations, a single model was used both with memory and problem-solving tasks.

The memory simulations showed that the model was able to capture several empirical phenomena in the way youngsters recall awalé positions, using assumptions that are the same as those used in the chess simulations. A key prediction from the chess simulations was that the capacity of visual STM should be three. A systematic analysis of performance as a function of STM capacity supported this prediction, although a capacity of four also seemed a reasonable estimate. With lower values, the model failed to reach the level of the average players, and with higher values, the model performed too rapidly at the level of strong players or higher.

The model was able to reproduce the main effects of skill and age. The simulations were weaker with respect to the number of holes occupied and the type of positions. In both cases, although the results were satisfactory in that the model reproduced the absolute value of most data points closely, the model made some incorrect predictions as well. With the number of holes occupied, it underestimated recall performance with four holes, and, with the type of positions, it underestimated the average players’ performance with double-threat positions. To explain the first anomaly, we speculate that players used some form of verbal coding with the positions containing four occupied holes, which would explain why the youngsters performed better than predicted by the model. As for the second anomaly, the dip shown by the model with the double-threat condition could be explained by the fact that the set of positions used for learning contains a small proportion of positions similar to the

<table>
<thead>
<tr>
<th>Criterion for Breaking Ties</th>
<th>Number of Chunks</th>
<th>Frequency of Association</th>
<th>Largest Chunk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tied</td>
<td>.510 (.003)</td>
<td>.207 (.004)</td>
<td>.348 (.003)</td>
</tr>
<tr>
<td>Only one</td>
<td>.361 (.002)</td>
<td>.150 (.003)</td>
<td>.348 (.003)</td>
</tr>
</tbody>
</table>

Note: The results are based on 500 replications. Learning was carried out using 7,559 positions, yielding networks with 11,249 nodes on average (SEM = 1.2). Testing was done on 8 positions. The average probability of finding the correct move by chance is .296 for this set of test positions. Values in parentheses indicate standard errors of the mean (SEM).
double-threat positions, perhaps because double threats are relatively rare in competitive
games played by strong players. Thus, average players had been more exposed to this type
of positions and were able to extract more chunks related to them.

Additional limitations must be noted with the memory simulations. Memorizing an awaél
position is a relatively simple task—at least compared with memorizing a chess position—as there are only 12 numbers to memorize. Thus, several strategies not related to the use of
chunks might be used, such as using verbal coding or employing mnemonics. In addition,
our strategy of keeping the model parsimonious had the consequence that the goodness of
fit was not high in all simulations. However, as argued earlier, our interest was more in
showing that the main theoretical assumptions of CHREST provide a reasonable explana-
tion of the cognitive process underpinning skill in awaél than in obtaining a close fit between
model and data by adjusting free parameters.

The simulations on decision making indicated that pattern recognition could explain at
least in part how awaél players find good moves rapidly, without analyzing many moves
ahead. There were also a few limitations in these simulations. The number of games used
for training the model was small. The consequence was that, except in the case of seen
positions, performance was not high. It is reasonable to assume that, as players are exposed
to a larger number of games, their behavior becomes similar to that observed with the simu-
lations where the positions had been seen before. This is particularly likely with endgame
positions, where the same patterns tend to recur often. In addition, it is likely that players
learn to associate not only moves to perceptual chunks, but also sequences of moves. Thus,
the model could be extended to carry out this kind of learning, as was done by the CHUMP
model for chess (Gobet & Jansen, 1994).

It must be noted that players also use additional information to visual chunks; this
includes known tactics and strategies, part of which is verbalizable, as can be seen in work-
shops where masters teach awaél. It is also clear that, at least in some critical situations,
awaél experts carry out a (limited) amount of look-ahead search. A final limit of the present
work is that selection among chunks was simplistic; in particular, some combination of cri-
teria could be used.

The model described in this article illustrates the potential role that computational mod-
eling can play in cross-cultural psychology. If a single theory can explain data from widely
differing tasks carried out in different cultures by individuals from those cultures, then one
can be confident that the same cognitive mechanisms are engaged. This in turn supports the
hypothesis that the basic cognitive mechanisms are universal (Kagitcibasi & Berry, 1989;
Lehman et al., 2004; N’Guessan, 1992; Retschitzki, 1990). The present study has shown
that a model based on the CHREST cognitive architecture can account for a number of
phenomena of awaél, related to both memory and problem solving, thus showing that the
postulated mechanisms and parameters—especially chunking based pattern recognition
and STM capacity limited to three items—are likely to underpin awaél skill, just as they
underpin chess skill. Although the current simulations could be improved in various ways
(e.g., by using more positions during learning, simulating more phenomena, and adding
search abilities and strategies to the model), this study makes a clear contribution to the
literature by being the first to show that the same cognitive architecture can be used cross-
culturally, and, more specifically, that chunking mechanisms play an essential role in a
game that is played in non-Western cultures. The study also adds yet another domain to the
increasingly long list of domains that have been simulated by CHREST, thus buttressing its claim to be a general cognitive architecture.

Notes

1. We do not consider that grammar-based models of natural language, which have been used with a variety of languages, invalidate this statement, as these models are limited to language and do not cover other aspects of cognition.
2. The capacity of CHREST’s visual STM has evolved over time. The first model used a capacity of seven items (Gobet, 1993a), but later versions used four (De Groot & Gobet, 1996b) and three (Gobet & Simon, 2000; Waters & Gobet, 2008) items.
3. The idea of anchor points gets some support from the terms used by players. For example, in Antigua, players call the leftmost hole the “foot.”
4. The average branching factor (number of legal moves in a position) is 3.5 moves in awalé, which is smaller than chess (35 moves) and Go (250 moves). Thus, the probability of finding the best move by chance is 1 / 3.5 on average, that is .286 (see also the notes of Tables 2, 3, and 4). Is awalé then too simple a game for studying cognition? We do not think so. Other complexity measures show that awalé is far from a trivial game. For example, whereas a maximum of two pieces can be involved in a move in chess and a maximum of one piece in Go, up to 48 pieces can be involved in a single move in awalé! See chapter 2 of Gobet et al. (2004) for details.
5. We did not use Problem 8, as this required a sequence of moves rather than the first best move.

References


**Fernand Gobet** is a professor of psychology at Brunel University (West London). In his research on expertise, he has investigated skilled individuals' abilities, focusing on perception, learning, memory, and problem solving. His theoretical work in computational modeling has led to the development of a computer model that puts together these psychological functions. This model has been used to simulate phenomena in chess and awalé expertise, problem solving in physics, concept formation, and the acquisition of language. He is coauthor of *Perception and Memory in Chess* (with A. D. de Groot) and *Moves in Mind* (with A. J. de Voogt and J. Retschitzki).

For reprints and permissions queries, please visit SAGE’s Web site at http://www.sagepub.com/journalsPermissions.nav.