Chess Masters' Hypothesis Testing in Games of Dynamic Equilibrium (Cowley, 2016)

Chess inculcates certain virtues such as foresight, patience, and the ability to accept the consequences of one's decisions.

—Benjamin Franklin (*US President*, 1786; cited in Hartston & Wason, 1983, p. 7)

A fundamental type of thought is the prediction of events (Craik, 1943). When people generate a hypothesis, they may consider the implication of the hypothesis assuming it is true, or they may attempt to predict the most likely outcome given some scenario (e.g., Fugelsang et al., 2004; Girotto, Legrenzi, & Rizzo, 1991). For example, people may anticipate a possible outcome when planning a future action or choice that they will make. This paper examines how people test their hypotheses when they are planning a course of action. Understanding how people test their hypotheses about plans of action may help in understanding how people commit errors, learn from past mistakes and develop better planning for the future (e.g., Roese & Olsen, 1995; Vygotsky, 1986).

This paper focuses on plans of action in chess players' hypothesis testing. Thinking in chess may consist of exploring different alternative hypotheses about what the best move is, in order to achieve a desired goal (Newell & Simon, 1972). Chess players must play moves that are so good that they may not be refuted (Saariluoma, 1995), implying that it is helpful to consider the possible alternative ways a plan may be falsified. In other words, how might a plan go wrong, for example, by an opponent responding to a plan in a way previously not anticipated? Chess presents an ideal paradigm in which to examine such questions because studies of chess have contributed substantially to the understanding of human problem solving (e.g., Newell & Simon, 1972; Gobet & Simon, 1996a; 1996b; 1996c), and chess provides a neat micro-world of competitive reasoning in which to investigate different aspects of thinking with unparalleled precision (DeGroot, 1965; Newell, 1990).

This paper examines how chess players anticipate an opponent's response to the moves they consider for play by suggesting that chess players evaluate moves they plan to play by searching for confirming evidence that their plan will work, or falsifying evidence that their plan will not work (Popper, 1963). This investigation of hypothesis testing in chess emphasizes a question about chess playing originally identified by DeGroot (1965), who published the first studies of chess during the 1940's (translated to English in the 1960's). How do chess players anticipate the ways a plan might be falsified in order to avoid making mistakes? The question has been neglected since and studies of chess have focused on other factors, such as memory for chess positions (e.g., Chase & Simon, 1973a; 1973b). DeGroot pointed out that the choice of a good move rests partly on foreseen possibilities for action *and* on the evaluation of their foreseen results (1965).

To address this question this chapter presents the first exploration of chess players' hypothesis testing. This paper provides one of the largest analyses of chess players' thinking for over thirty years (DeGroot, 1965; Newell & Simon, 1972). Protocols were selected from ten players who thought aloud while choosing moves for play for six positions each. Sixty protocols were analysed (minus three that were inaudible) and experts were compared to non-experts (5 master level players and 5 novices). The protocols were selected from a larger, previously unanalysed thesis set of data (Cowley, 2002). The data were analysed by mapping out all move sequences considered by each player for each position, and by using one of the most powerful chess programs in the world to objectively evaluate the outcomes of move choices (*Fritz 8*). The central prediction was that expert chess players would attempt to falsify, by anticipating opponent moves that would ruin their plans more readily than novices,

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

because experts may be better at detecting mistakes and recovering from them than novices.

In what follows this paper will describe in detail why chess presents us with an ideal domain in which to investigate hypothesis testing, by outlining the methodological advantages of the chess domain. Two main sorts of theories of chess expertise are described; chunking theories and search theories, and a new theoretical component of chess expertise based on a hypothesis testing framework derived from the experiments reported in this paper are reported. An experimental analysis of chess expertise is outlined and the implications of a hypothesis testing theoretical component of chess expertise for cognitive theories of how people test hypotheses are discussed.

Methodological Advantages of Chess Experiments

Studies of chess have contributed substantially to understanding human cognition, including how people solve problems (Newell & Simon, 1972), how people chunk knowledge together to remember it (Chase & Simon, 1973a; 1973b), and how people develop their expertise (DeGroot, 1965). Chess findings provide strong external validity; results observed in chess experiments have contributed to explanations of expertise in non-game domains such as physics (Larkin, McDermott, Simon, & Simon, 1980; Chi, Glaser, & Rees, 1982), computer programming (McKeithen, Reitman, Rueter, & Hirtle, 1981), and music (Sloboda, 1976). Accordingly, the findings concerning chess players' hypothesis testing, and chess masters' hypothesis testing in particular, may lead to the generation of similar explanations of expert thinking in non-game domains, such as scientific and political thinking (e.g, Fugelsang et al., 2004). The findings may also have implications for understanding thinking in domains analogous to chess where interaction with a collaborator or an opponent is required, for example, in expert military strategy (Mallie, 2001).

Another advantage is that expertise in chess can be defined objectively and categorized relative to other experts by an established standard rating system. This rating system ensures that we have a relatively accurate system for classifying levels of expertise, more so than in some domains where it is difficult to classify expertise, such as in leadership or creative domains (e.g., Gardner, 1983; Charness, 1991). For example, the Elo rating system in chess classifies the level of expertise that players have obtained (Elo, 1978). Elo systems calculate an expected playing strength rating on the basis of competitive tournament games. The stronger the player one can defeat the more points one's rating is increased, and the weaker the player one is defeated by the more points one's rating is decreased. Ratings vary between approximately 1000 for an absolute novice and over 2800 for the world chess champion.

A further advantage is that chess has a 'unique' notation for describing what a player is currently representing while thinking about what move to play. Each player is fluent in this notation, and can verbalise what they are thinking about, with this verbalisation having minimal interference with thought processes (Ericsson & Simon, 1993). The availability of this notation allows an experimenter who is also fluent in it to transcribe a player's thinking aloud, and map each and every move a player verbalises. The notation is called *algebraic notation* and it is described in detail in the next section.

Algebraic Notation in Chess

Chess is a board game where the overall goal is to move your pieces around in a purposeful manner, in order to checkmate the opponent by attacking the opponent's king piece, and eliminating all the possible ways the opponent king can escape your attack. Chess thinking may consist of exploring different alternative paths or 'moves' in a 'problem space' (Newell & Simon, 1972). The problem space consists of the initial problem state, that is, the start of the game; intermediate problem states, for example, capturing an opponent piece, and the end

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

state, checkmate. Progress from state to state is achieved through operators and in chess the operators are the ways the chess pieces are allowed to move. For example, a bishop operates diagonally backwards and forwards and captures on the square on which it lands for any one move. At the beginning of a game of chess the two players have equal numbers of pieces and theoretically almost equal chances of a win. Algebraic notation means that chess operators can be formalized mathematically and players' verbalizations of their thinking can be recorded in the common short-hand code used to refer to moves. Consider a simple example of algebraic notation with reference to Figure 1 below.
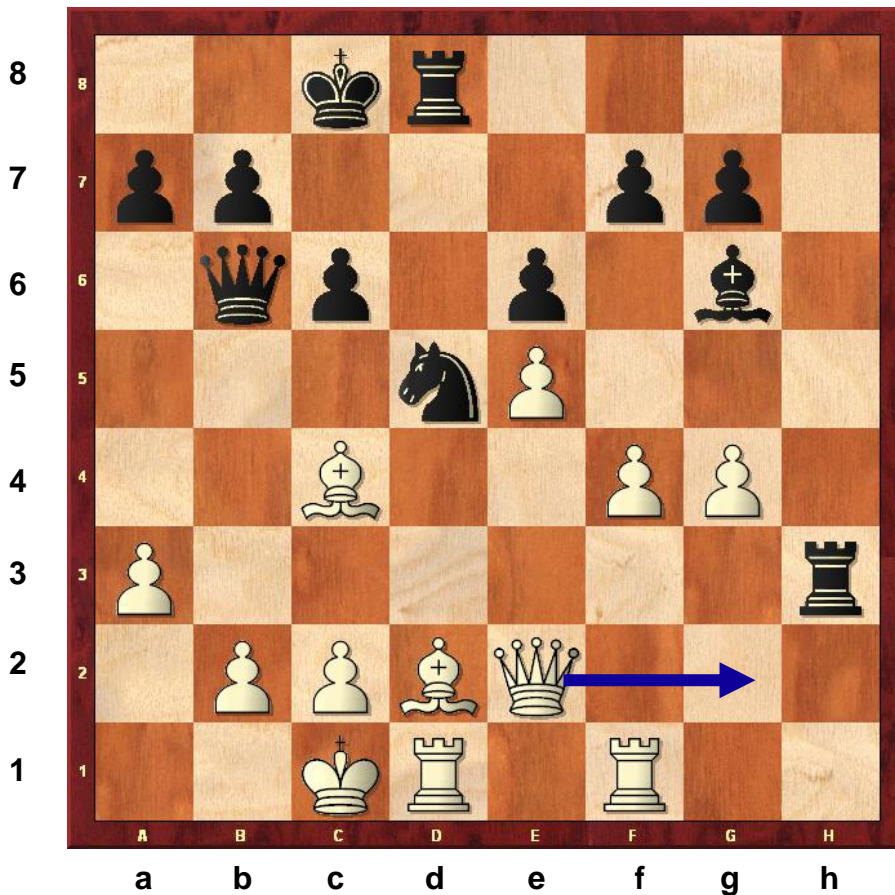

Figure 1: A representation of a chess board middle game, in which it is white to play.

Suppose a player faced with this board is invited to think-aloud. The player says "If I play Qg2", which means the player is thinking about moving the queen (Q) to the square called g2. The arrow indicates that the square that the queen is perceived to move to. Each square on the chessboard has an algebraic coordinate. Starting from the left hand corner and moving horizontally to the right, each square (x8) is labelled with a letter as Figure 1 shows. The letters range from a - h. Again starting from the left hand corner but this time moving vertically upwards, each square (x8) is labelled with a number. The numbers range from 1 – 8. So, the coordinate of the first square in the left-hand corner is known as a1. The square to a1's immediate right is b1 and so on. There are sixty-four coordinates altogether. The algebraic notation provides chess players with the means of recording where pieces have moved in a game of chess. The pieces have their own symbols. The symbols are King = K; Queen = Q; Rook = R; Knight = N; Bishop =B; Pawn moves are recorded merely as coordinates (e.g. b4 = a pawn has moved to the b4 square). For example a move could be Ne4

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

(the knight has moved to the e4 square). Typically chess players do not verbalise one move in isolation, such as Ne4 or Qg2. They often generate a sequence of moves of the form "If I play Qg2, then you might play Nd7, and if you play Nd7 then I'll play…". Next we turn to the most influential experimental studies of chess players' thinking, which used this notation. First, we will describe the theoretical debate about what constitutes chess expertise: the ability to search through more moves than one's opponent to find a good move, or the ability to reproduce good moves from large repositories of chess knowledge built up with extensive amounts of practice. This paper will describe an experimental analysis of hypothesis testing in chess playing which helps to resolve this debate. The exploratory analysis described may elicit a new hypothesis testing component of chess expertise and outline the implications it may have for cognitive theories of expertise, human learning, and reasoning more generally.

Theories of Chess Expertise
The main theoretical debate that exists for chess reasoning is between chunking and search theories. Each theory type is explicated below:

Chunking Theories
Chunking theories of chess expertise focus on how experts store their knowledge of a domain, that is, how smaller related pieces of knowledge can be grouped together into larger pieces of knowledge called chunks (Chase & Simon, 1973a; 1973b). The chunking theory of expertise was developed initially from De Groot's early examination of 22 chess players (6 grandmasters, 4 masters, 7 experts and 5 strong club players). The players were requested to think-aloud while deciding what move they would play next (DeGroot, 1965). Their protocols showed that they considered the same number of moves, and searched to the same depth in move sequences. (Search depth refers to how many moves ahead a player thinks about from the current position the player is thinking about. A player who thinks about what move he or she will play and what move the opponent will play in response, and then what move he or she will play in response to that, has searched to a depth of three moves. Each move is usually called one ply, and so the player's search depth is 3 ply.) But when players were shown board positions for two to fifteen seconds, and asked to recall where the pieces had been, the masters' recall was better (93%) than the recall of the experts and the strong players (72% and 51%). The result has been replicated (Chase & Simon, 1973a; 1973b; Gobet, de Voogt, & Retschitzki, 2004). It has been taken to indicate that the masters' greater playing experience leads to greater familiarity with chess configurations likely to occur in a game (DeGroot, 1965).

In a study of three chess players (a master, a class A expert - which is a strong club player with a US rating between 1900 and 2100, and a beginner), the participants were presented with normal chess board positions and random positions (Chase & Simon, 1973a; 1973b). In the random positions, the pieces were scrambled to achieve a random pattern unlike any that would have occurred in a real chess game. Each of the three players showed poor recall for the random positions, worse even than the beginner's recall for normal positions (less than 17%). The result corroborates the idea that experts may be better at remembering information pertinent to their domain of expertise because they have experience in that domain. Moreover, experts 'chunk' information more effectively than novices (Chase & Simon, 1973a; 1973b). In one task, two chessboards were placed side by side. Positions were set up on the left board and the 3 players were required to reconstruct the pieces on the empty chessboard on the right. They were allowed to glance back and forth between the position and the empty board and they were videotaped. When players placed pieces together on the empty board within a glance (without looking back at the position), it was taken to indicate

that these pieces had been 'chunked' together in memory. Two pieces placed one after another rarely exceeded two seconds. A glance was assumed to establish the boundary of a chunk and pieces placed within a limit of two seconds were assumed to belong to the same chunk. Chunks were found to be related to one another in a number of different ways including spatial proximity, colour, similarity, attack or defense relations, pawn chains, and castled king formations. The average size of chunks in recall correlated with the 3 players' level of skill. The master placed 2.5 pieces on average within a glance, the class A expert 2.1 pieces, and the beginner 1.9 pieces. The master also encoded the chunks in shorter glances suggesting that experts may be faster encoders of domain specific knowledge. Accordingly, the ability of experts to recall briefly presented board positions may rely on better pattern recognition (more information encoded within a glance and faster encoding times). Further support for this chunking model of expertise was provided by an early computational model called MAPP (Memory Aided Pattern Perceiver) (Simon & Gilmartin, 1973). The program held what was considered to be a large number of chunks at the time and it encoded board positions into its short-term memory by recognizing chunked patterns. Two versions of the program were generated and the one with more chunked patterns performed better (1114 versus 894 patterns). According to the chunking theory each of these patterns can quickly be identified and can be accessed in a players' long-term memory through a label (Frensch and Sternberg, 1991, p.356). This label matching process activates retrieval of information about that chunk (Saariluoma, 1995). According to Chase and Simon (1973) chunks reduce the cognitive load in working memory. Since beginners do not typically have access to a large store of cognitive chunks they must try to recall the 24 individual pieces and their locations. The brief presentation of five seconds assumes that only short-term or working memory is being used, and typically it has been found to have a capacity of 7 items plus or minus 2 (Miller, 1956; Baddeley, 1999). Therefore beginners find it difficult to store up to 24 pieces in working memory, whereas masters can store a chunk containing, for example, 3 pieces as one item in working memory, increasing their working memory capacity for briefly presented positions. In random positions there are no chunks available for recognition and so the skilled players' performance is no better than novices (Chase and Simon, 1973).

In fact, experts may be able to access a considerable amount of information, more than an isolated chunk (Gobet, 1996). Board position knowledge may be stored in schematic retrieval structures or *templates,* that is, an abstraction of a set of similar positions and general plans and possible moves that follow from them (see Gobet, 1998; Gobet et al., 2004). Templates are more general than any specific type of position. Support for the view comes from studies of chess masters' memory. Gobet & Simon (2000; for a review see Gobet et al., 2004) compared five masters with eighteen experts and seven class A players. The players reconstructed a rapidly presented board position within a two second interval. Masters remembered and reconstructed the locations of more pieces from rapidly presented board positions than the experts and class A players. In addition, the sizes of the largest chunks the masters remembered were larger than the others (15-16 for masters, 9-10 for experts, 4-5 for class A players). For example, a chess player may be familiar with an opening known as the 'Sicilian Dragon' in which the pawn formation is shaped like the profile of a Chinese dragon spanning the width of the board (Gobet, de Voogt, & Retschitzki, 2004; Nunn, 1999). Their template may contain information about the general structure of the positions that arise from this opening and general plans that follow on from achieved positions. A master who is familiar with the Sicilian Dragon opening would be able to recall a briefly presented Sicilian Dragon position in chunks that contain more than 3 pieces. For example the seven pawns in the pawn structure can be encoded in a single chunk. A position template is structurally different to a set of chunks because of its integration of an entire

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

position. Knowledge structures may become more integrated with increasing levels of expertise.

Accordingly the chunking and template theories predict that experts rely on domain knowledge to deal with routine problems with which they are familiar. These theories provide convincing explanations for how experts choose a good move, that is, a move that improves their position and does not lead to a worsening of their position. But why do experts sometimes make a bad move, that is, a move that worsens their position? Expertise is often associated with a lack of error but experts do make mistakes (Green & Gilhooly, 1992). Experts have not always been experts and they continue to have experience of making mistakes, as losses at the highest levels of competition testify. Perhaps chunking and template theories are better at predicting performance for routine problems, which experts are familiar with. What happens when experts are confronted with a problem that does not map neatly onto their knowledge?

Search Theories of Chess Expertise

In an early investigation of how experts deal with novel problems, 3 experts and 3 novices were asked to choose moves in randomized chess positions, where the pieces had been scattered about the board, until few regularities of chess configurations were present (Holding & Reynolds, 1982). The experts chose better quality moves for play. How did the experts choose better moves for play if they could not use their superior knowledge? The experts may be able to use their superior search capabilities to choose better moves (Holding & Reynolds, 1982; Holding, 1985). The implication was that deeper search may explain how masters anticipate their opponents' moves, prepare a counter-move in advance, and thereby choose better quality moves. Experts have been found to search a larger problem space, that is, they consider more moves than novices, when the board positions are more complex (Gobet & Campitelli, 2002). But experts have only been found to search deeper than novices when the problems they are presented with are not routine and do not demand reproducing solutions already known.

Overall reproductive-based theories, such as the chunking and template theories appear to explain more of the data on experts' skill than search-based theories. The overwhelming support for the view comes from evidence that experts' first moves (when no search has taken place) are of a consistent, 'good enough' quality. Klein, Wolf, Militello & Zsambok (1995) found that the protocols generated by 8 highly skilled players (with a rating range between 1700 to 2150) contained first verbalized moves that were good enough for play as rated by an independent grandmaster. Moreover in 'blitz' chess, in which each player is allotted five or ten minutes for an entire game, moves must be made quickly. Players spend about an average of 5 seconds per move in blitz chess compared to an average of 180 seconds per move in standard chess. Despite very little time for extensive search processes, performance in blitz chess shares 81% variance with players' ratings based on their performance in standard games of chess (Burns, 2004; Chabris & Hearst, 2003). Likewise, the performance of world chess champion Kasparov did not decrease when he played a simultaneous exhibition against a team of strong chess experts (the Israeli national team with players of grandmaster strength) in which he had little time between choosing moves for different positions (Gobet & Simon, 1996a; 1996b; 1996c). Perhaps experts may search more deeply than novices only to solve problems with which they are unfamiliar or which are novel in some way.

But how informative is depth of search as a measure of the key difference between experts and novices? Consider players thinking about a move and thinking about the move their opponent may make in response. There may be many possible moves the opponent

could make. One player may think immediately about the possible move the opponent could make that would *refute* the player's plan. The player's search is just two moves deep (their own and their opponent's move), but it reveals a crucial refutation of their plan. Another player may instead fail to think immediately about the refuting move of the opponent. Instead the player may think of some of the possible moves the opponent could play that will not refute the player's move, and may even lead to an advantage for the player. The player's search may be deep but confined to opponent moves that *confirm* their plan. A third player may think of just one or two moves the opponent could make, and each may be moves that would fit with the player's plan. The player's search is not deep and it is confined to confirming moves. A quantitative comparison of the depth of search based on the surface structure is not informative: the first player searched to just two moves deep (but found a potential refutation), the second player searched four or five moves deep (but found no refutation), and the third player searched to the same depth as the first player (but found no refutation). An important difference may be present in the deeper structure of the search. An expert may think about the strongest move an opponent can play - a refuting move, and so reduce the need to examine a large search space. An expert may be distinguished more by the selection of *what* to investigate than by the depth of calculation (DeGroot, 1965; Kotov, 1971).

Hypothesis Testing in Chess Expertise

The aim in this paper is to examine how chess experts err, and in particular how they detect error and recover from it, a question never before addressed. There are hints in the literature that expert chess playing requires more than extensive and accurate memory retrieval, as experts do not automatically accept the first move they generate when they have time. Instead they appear to search through a problem space as revealed by their think-aloud protocols (DeGroot, 1965; Newell & Simon, 1972). Why do experts focus their efforts on search given that their first moves generated from their stored knowledge tend to be good enough (Gobet & Campitelli, 2002)? This paper proposes that expert knowledge may not be perfect, or that the knowledge experts retrieve from long-term memory may not match the current position best. To avoid choosing moves that lead to error as a result of inaccurate retrieval, chess experts may evaluate the moves they consider, and their evaluation can usefully be conceptualized as a form of hypothesis testing. This paper predicts that expert chess players may tend to try to detect and avoid error by engaging in hypothesis testing, that is, they may attempt to falsify their move plan, especially when they have time to check the accuracy of their thinking. In other words expert knowledge may help people to falsify *their own* hypotheses. The account implies that the purpose of expert search is to detect where an opponent can refute a move planned for play, to identify errors. Choosing a move in chess may depend on accessing a large repository of domain knowledge about possible moves for play, including possible opponent moves. All players may try to think about their opponent's response to their move, but a key difference between expert and novice search may lie in the anticipation of how opponent moves can have negative, rather than positive, consequences for a planned move.

DeGroot (1965) suggested four stages to choosing a move in a game of chess, corresponding to a *progressive-deepening* strategy (De Groot, 1965). The four stages are orientation, exploration, elaboration, and proof. During the orientation phase the player adjusts to the position, weighs up the number and placement of pieces for the player and the opponent, and identifies strengths and weaknesses of the position. The function of the phase is the formation of a specific problem conception that may include a selective set of possible board plans that correspond to a set of considered moves each aimed at attaining a specific

board goal. There may be a preference or 'favourite' within this set of moves. These plans, considered moves and anticipated solution are partly tentative and hypothetical. The player's conception of the board problem is a 'working hypothesis' (De Groot, 1965, p. 395). During the second phase of exploration, the player begins to explore possibilities in the problem space. During the third phase of elaboration, the player continues the exploration in more detail and may repeat some investigations. Finally, during the proof phase, the investigations are evaluated.

This paper conceptualizes each move a chess player considers to be a hypothesis, and the opponent moves they consider in response to each move as tests of the hypothesis. These tests may be either confirming (the opponent's move fits in with the player's plan, that is, it leads to an improvement in the player's position), or they may be falsifying (the opponent's move refutes the player's plan, that is, it leads to a worsening of the player's position). The paper examines experimentally expert and novice chess players' evaluations of the moves they considered in response to normal and random board positions. Evaluation is conceptualized as a hypothesis testing process.

Evaluation of Chess Moves by Experts and Novices
An experiment was designed to test whether there were differences in the types of evaluations generated by experts and novices when they thought about what move they would like to play in a given position. The aims of the experiment and its method are outlined. The results from a previously reported thesis retrospective evaluation analysis is summarised in the experiment (i.e., Cowley, 2002). Then the first report of the previously unanalyzed protocol data collected in the experiment is outlined.

I designed a experiment to examine how hypothesis testing could be conceptualized as the evaluation of moves considered for play by chess players (Cowley & Byrne, 2004, for details). The experiment tested ten experts and ten novices (including the full population of chess masters living in Ireland at the time). The aim of the experiment was to measure a) retrospective evaluations of chosen moves, and b) think-aloud protocols when choosing moves. The experiment and the results for the retrospective evaluations are reported in Cowley (2002) as an unpublished BA research project, and while we cannot report those data on retrospective evaluations in full here, we will outline the main results and methodology. The results showed that expert chess players (including master and non-master level experts) chose a higher quality move for play than novices at the end of their thinking time in the normal positions. While experts tended to retrospectively evaluate a greater proportion of the move sequences they thought about as leading to a negative outcome, novices tended to retrospectively evaluate a greater proportion of the move sequences they thought about as leading to a positive outcome. This result suggests that experts chose higher quality moves for play because they tended to anticipate the opponent moves that led to a negative outcome for their board position. Novices may have chosen lower quality moves for play because they tended to anticipate how their moves could lead to a positive outcome even when they objectively led to a negative outcome. Experts thought about how their plans could be falsified, but novices thought only about how their plans could be confirmed.

The focus in this paper is to examine this result in greater detail by analysing the previously unanalysed and unreported protocol analyses. First I summarise the methodology of the retrospective evaluations experiment reported in Cowley (2002) in some detail given that the protocol analysis is based on a subset of the participants who carried out the retrospective evaluations, and the materials and design were the same. Then I will summarise its methodology, and describe how the result led to the protocol analysis reported for the first time in this paper.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Retrospective evaluation

In the earlier retrospective evaluations experiment reported in Cowley (2002), there were 20 participants. The 20 participants (19 men and 1 woman) were registered members of the Irish Chess Union. They were experienced novices (mean rating of
1509) and experts (mean rating 2240). The expert group included experts from different Elo categories of expertise, including one grandmaster (Irish Elo >2500) two international masters (Irish Elo > 2300), three Fide masters (Irish Elo > 2200, i.e. International Chess Federation masters), and four initial category experts (Irish Elo >
2000).

The materials were six board positions, three normal and three random (as well as an initial practice position). The board positions were chosen from games in chess periodicals. They were middle game positions with 22-26 pieces to ensure complexity and to rule out the chances that the masters had seen them before. The positions were chosen with the assistance of a chess expert (who was not a participant in the study). Importantly, they were 'equality outcome' positions, where there were equal chances with best play for both black and white pieces (see Appendix A). This constraint ensured that there would be no obvious confirming or falsifying move sequences. These properties were selected to ensure complexity and to rule out the likelihood that the masters would have seen them before. Although they were equality outcome positions they were not 'quiet' positions such as those where very few possible moves are available to the players and they inevitably lead to a draw. Instead, they were selected to be in the style of De Groot's (1965) position A where the position is 'lively' with many possible alternative moves to choose from. Any colour preference of player's to play with white or black was controlled by balancing the number of times they played the white or black side of the positions.

Random board positions also contained twenty-two to twenty-six pieces but the pieces had been scattered. They were randomized until positions were achieved to satisfy the constraints that neither king was in check, no pawn occupied the first or eighth rank, and they were equality outcome with best play, that is, they were positions of dynamic equilibrium. They were also designed and checked by an independent chess expert and the chess program *Fritz*. These random board positions with equality outcomes were designed specifically for this experiment. The first two constraints make it possible to play and choose a move in randomized positions (Holding & Reynolds, 1982). See Figure 1 for an example of a chess position used (position 1).

The participants' task was to, "choose a move you would play in the way you are used to going about choosing a move in a real game". They were given instructions to think-aloud, and their verbalizations were recorded by dictaphone (See Appendix E for the experimenter's think-aloud script). Moves examined by the player during think-aloud are verbalized using algebraic chess notation. These moves were recorded by the author who is fluent in algebraic notation (See Appendix F for the experimenter's recording sheet).

Three minutes thinking time was allotted for choosing a move as it is just over the average time per move in tournament play. Exposure for each board position was timed using a standard tournament chess clock, each clock was set at three minutes and when the clock's flag fell participants were told that their time was up. To accurately access hypothesis testing we also needed participants to provide us with an evaluation of each move sequence that they examined. However, spontaneous evaluation in chess has a low probability of verbalization (Newell & Simon, 1972). Accordingly we used a combined methodology of think-aloud followed by retrospective evaluation (See Appendix E for the experimenter's retrospective evaluation script). Verbalized move sequences were recorded not only by dictaphone, but also by the experimenter (the author) in algebraic notation concurrent with think-aloud. The

experimenter asked the participants for their evaluation of each move sequence, by first saying back the move sequence immediately after each chess problem to reduce retrospective error and interference (Ericsson & Simon, 1993). The participants were then asked to evaluate each move sequence as having led to a positive, negative or neutral outcome for their positions. When players evaluated a move sequence as leading to a positive outcome it was scored as confirmation. When players evaluated a move sequence as leading to a negative outcome it was scored as falsification, and when players evaluated a move sequence as leading to a neutral outcome it was scored as neutral.

The findings showed that novice chess players evaluated their move sequences as leading to a positive outcome more readily than experts, even though they chose lower quality moves to play than the experts at the end of their thinking time. Experts evaluated their move sequences as leading to a negative outcome more readily than novices, and they chose higher quality moves to play than the novices. The results suggest that experts could falsify their hypothesized moves by evaluating them as leading to a negative outcome, and then they chose higher quality moves than the novices. Experts can falsify their own hypotheses to abandon low-quality hypotheses (Popper, 1959).

However, the players' retrospective evaluations are subjective, for example, a player may evaluate a move sequence as leading to a positive outcome but objectively it leads to a negative outcome. The player may have anticipated only how opponent counter-moves to their chosen move lead to a more positive outcome for them, and they may not have anticipated how opponent counter-moves would lead to a negative outcome. Alternatively the player may have anticipated opponent counter-moves that would lead to a negative outcome, but they may have misinterpreted them as leading to a positive outcome.

The protocol analysis reported in this paper aims to discover whether players' subjective evaluations corresponded to objective evaluations. One of the most powerful chess programs was used to achieve objective evaluations of the verbalized chess moves (*Fritz 8*). First, the players' think-aloud protocols were transcribed, and then the problem behaviour graphs were created based on the transcriptions. Each possible move mentioned by each chess player was mapped out to create a problem behaviour graph, and the properties of the problem behaviour graphs were analysed to discern the role of mental representation in chess players' hypothesis testing. We now turn to the experiment.

A protocol analysis of hypothesis testing by masters and novices
The experimental analysis of the think-aloud protocols reported here is new and previously unreported data. A subset of the protocols were selected including the protocols collected in the retrospective evaluation experiment that was summarised in the previous section. A new and more complete experimental and theoretical analysis of chess players' hypothesis testing was explicated. It is instructive to focus on the master level players (for comparison with masters studied in the chess literature previously) and to this end the think-aloud protocols of five master level players (i.e. 1 Grandmaster, 2 International Masters, and 2 Fide International Chess Federation Masters) who took part in the earlier experiment were selected, and compared to the think-aloud protocols of five novice chess players, chosen at random from the full sample of novices. First the aims of the protocol analysis and its method are described. Then the first report of these data is provided.

Objective evaluation of chess moves: Fritz 8
Fritz 8 is a chess program. It plays chess to the standard of the world's current top five grandmaster level players, and at its conception drew in a match with the world chess champion Vladimir Kramnik (Bahrain, 2003). In analysis mode, Fritz 8 can produce an

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

evaluation of a position after a chess player plays a move. The evaluation tells a player whether a move will objectively lead to a positive outcome, a negative outcome, or a neutral outcome regardless of what the player may subjectively wish the outcome to be (see Appendix B for the setup functions in the use of the program *Fritz 8*). In fact as Table 1 illustrates, there are nine types of hypothesis test in chess based on whether a subjective evaluation of a move leads to a positive, negative or neutral outcome; and whether the evaluation of the move objectively leads to a positive, negative or neutral outcome.

Table 1: The nine possible hypothesis types based on the subjective and objective evaluations of move sequences

Retrospective evaluation by chess player	Objective evaluation by Fritz		
	Positive (+)	Negative (-)	Neutral (=)
Positive (+)	+/+	+/-	+/=
Negative (-)	-/+	-/-	-/=
Neutral (=)	=/+	=/-	=/=

Key: '+' refers to a positive evaluation, '-' to a negative one, '+/-' means the player's evaluation was positive and the program's evaluation was negative.

As Table 1 shows there are two types of *confirmation* in chess, (i) objective confirmation in which a player evaluates a move sequence as leading to a positive outcome and objectively it leads to a positive outcome ('+/+' in Table 1), and (ii) confirmation bias in which a player evaluates a move sequence as leading to a positive outcome and objectively it leads to a negative outcome ('+/-'). There are also two types of *falsification* in chess, (i) objective falsification in which a player evaluates a move sequence as leading to a negative outcome and objectively it leads to a negative outcome. In this case the player has anticipated the opponent counter-moves to his chosen move that lead to a negative outcome ('-/-' in Table 1), and (ii) falsification bias, in which a player evaluates a move sequence as leading to a negative outcome but it objectively leads to a positive outcome for them ('-/+'). They may not have anticipated how opponent counter-moves would lead to a more positive outcome for them. Finally there are two types of *neutral evaluation* in chess, (i) objective neutral evaluation in which a player evaluates a move sequence as leading to neither a positive nor a negative outcome and objectively it leads to this outcome ('=/='); (ii) neutral bias in which a player evaluates a move sequence as leading to neither a positive nor a negative outcome. However, the player may not have anticipated how opponent counter-moves would lead to a more positive outcome or a more negative outcome ('=/+' or '=/-'). Players may also evaluate a move sequence as positive when it is in fact neutral, a positive bias similar to confirmation bias ('+/=') or they may evaluate it as negative when it is in fact neutral, a negative bias similar to falsification bias ('-/='). Neither of these evaluations leads to error in chess however, and so we do not include them in the initial analysis of confirmation and falsification biases, although we do include them in a second broader analysis of positive and negative testing.

Protocol analysis in chess and the creation of problem behaviour graphs
Our aim in the experimental analysis we report here is to map each possible move mentioned by each chess player so that their move sequences can be inputted into the chess program

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Fritz 8. We wished to compare the *Fritz 8* objective evaluation with the subjective evaluation generated by the player to calculate the amount of objective falsification (-/-), and the amount of biased confirmation (+/-) etc. Deriving the objective measures from *Fritz 8* requires a detailed protocol analysis of think-aloud scripts. Protocol analyses of think-aloud scripts are time-consuming. Studies of chess have tended to rely instead on accuracy and time to recall board positions, or quality of final move choice (e.g., Holding and Reynolds, 1982). Studies that have used protocol analysis in problem solving tend to rely on very small sample sizes of just a few participants (e.g., Newell and Simon, 1972), although one chess study included a protocol analysis of five of the world's top grandmasters at that time (De Groot, 1965). We wished to carry out a protocol analysis on a reasonably large sample of players and so we selected a subset of the think-aloud scripts provided by ten of the players in the experiment reported in Cowley (2002), which reports the retrospective evaluations but not the protocol analysis.

First, the aim was to transcribe the taped think-aloud protocols for ten players, for six positions each (three of the positions were normal chess board positions, and three were random board positions). The transcribed think-aloud protocols were *segmented* into episodes, move by move (see Appendix C for all of the segmented protocols). Fifty-seven (3 protocols were inaudible including one FM normal board position and two GM random board positions) *problem behaviour graphs* were constructed using Newell and Simon's guidelines for the normal chess board positions (see Appendix D for a subset of the problem behaviour graphs generated by the experimenter).

To illustrate we present a small section of a master's problem behaviour graph in Figure 2, and we explain each corresponding move in the text that follows.
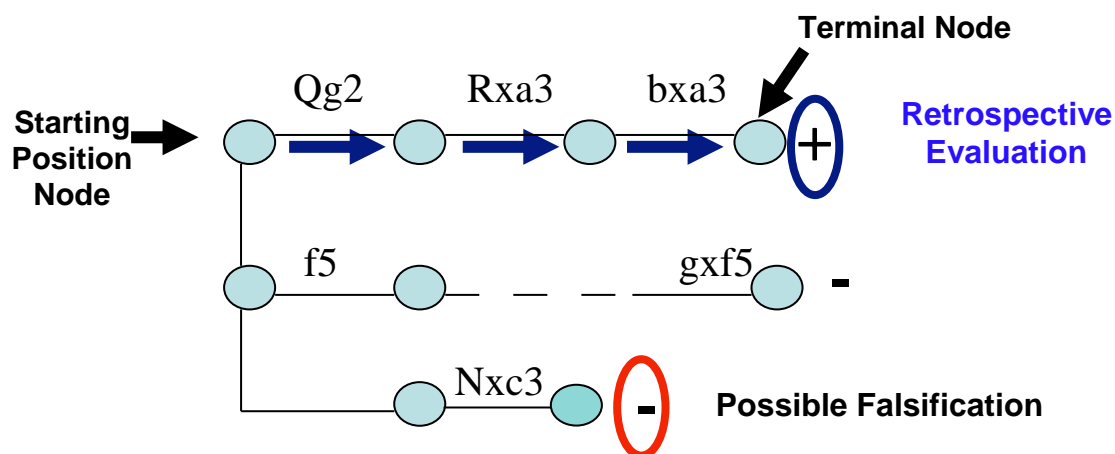

Figure 2: A section of a problem behaviour graph constructed from the chess master's protocol.

Each line across represents a move sequence. The order of search is from left to right, then down. Each circle (i.e. node) represents a new position following a move made in the problem space. For example, a player says Qg2 and it means the player is describing the possibility of moving his queen to the g2 square. Then the player says Rxa3 which refers to the player describing this move as a possible reply from the opponent to his Qg2 move (The opponents rook moves to the a3 square, and the x refers to the fact that the rook captures a piece, in this case a pawn). The next move bxa3 means the player says his pawn on the b-file can move to a3 to capture the opponent's rook piece. The plus sign shows that the player evaluated this move sequence as positive for the player, in their retrospective evaluation. The

minus shows a negative evaluation. The dashed line indicates that moves have not been specified for either the self or opponent. We label these *skipped* moves. For example, in the second move sequence, the participant first says f5 as a move for the player which means a pawn moves to the f5 square (note that pawn moves do not have the p in front of them). Then the player says gxf5 which means the pawn on the g file of the board captures a pawn on the square f5. This second move is only possible for the player and not for the opponent. Hence, the player has generated a move sequence describing only his own moves and has omitted any description of opponent moves. As the problem behaviour graph shows, we can incorporate the retrospective evaluation (positive, negative, or neutral) with the think-aloud move sequences. A subset of the problem behaviour graphs created for masters and novices think-aloud protocols of normal and random positions are presented in Appendix D.

To secure objective evaluations, we then used *Fritz 8* to evaluate the chess position occurring at the final move (i.e. terminal node) of each sequence (each line). Use of *Fritz 8* was essential to enable us to identify move sequences that would genuinely be positive for a player, and to discriminate them from move sequences that a player identified as positive for them but which could ultimately end negatively for them if played. Thus confirmation bias could be identified as a move sequence that was evaluated as leading to a positive outcome when it in fact leads to a negative outcome. Likewise, we were able to ensure that move sequences evaluated by a participant as negative were objectively instances of falsification. In other words, we could discriminate clearly between the nine types of hypothesis test presented in Table 1.

The problem behaviour graphs allowed us to measure several properties of chess masters' hypothesis testing. The first property of interest is the sorts of hypothesis tests that chess players generate. We wished to identify the confirming and falsifying hypothesis tests. A second property of interest is the quality of the first verbalised move. A third property of interest is the number of individual moves. A fourth property of interest is the search depth and the fifth property of interest is the search breadth (we define these measures shortly). The sixth property of interest is the different sorts of articulated move sequence: complete, skipped moves, base skip move, and ambiguous. The final property of interest is the point at which skipped moves occur.

For normal board positions we predicted that chess masters would use falsification to detect moves they examined for play to which an opponent could reply with a refutation (the -/- cell). Domain knowledge was predicted to facilitate hypothesis falsification, and expert hypothesis testing predicted to proceed by seeking falsifying instances of a hypothesis (see DeGroot, 1965; Gobet, 1998; Gobet et al., 2004). Access to expert knowledge may help chess masters to consider a larger number of possible alternative moves for both themselves and opponents than novices (see Johnson-Laird & Byrne, 1991). Further, chess masters may be better at representing their opponent's moves explicitly than novices; they are better than novices at detecting an opponent move that falsifies a hypothesized plan. Cognitive expertise may help chess masters to consider more possible alternative moves because they can represent their hypothesis testing more easily due to many years of practice (e.g., Ericsson & Kintsch, 1995). Novices were predicted to exhibit more evidence of confirmation bias suggestive of how people have been found to test their hypotheses in the standard hypothesis testing literature (e.g., Wason, 1960; Tweney et al., 1980; Mynatt, Doherty, & Tweney, 1978; 1979; Mynatt, Doherty, & Dragan, 1993; Poletiek, 1996). Novices were predicted to be less able to detect moves to which an opponent could reply with a refutation, and only see how these moves could lead to good outcomes. For random board positions, where the pieces have been scrambled about the board coordinates in order not to match expert knowledge, masters were predicted to no longer be able to falsify. They would perform similarly to novices

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

because their expert knowledge would no longer be able to help them to falsify their own hypotheses.

Method
Participants
Protocols from 10 of the 20 chess players who took part in the experiment reported in Cowley (2002) were chosen for analysis. 5 master level players were selected: 1 Grandmaster, 2 International Masters, and 2 Fide Masters. The selected experts conformed to the following criteria: they had an international master title of Fide Master, International Master, or Grandmaster; they had represented Ireland at international level at a Chess Olympiad team event (Chess Olympiads are the equivalent of Olympic games for chess, and are independently run by the world governing chess federation, *FIDE*); they had an ICU rating above 2250, and they were active players at the time of the study. They were compared to 5 novice chess players chosen at random from the sample of ten novices reported in Cowley (2002). The minimum and maximum ratings for this sample of novices were 1265 and 1511 respectively (M = 1413).

Materials, design and procedure
The between participants factor was the level of expertise (master or novice), and the within participants factor was the type of board position (normal or random) and the design was a 2x2 mixed design. Participants were given six board positions, three normal and three random positions as described earlier (see Appendix A for the six position diagrams).

Each participant received three normal board positions and three random board positions. The board positions were presented in randomized order. Each participant was also presented with a practice position to help them become familiar with thinking aloud before they moved on to the six experimental board positions.

Protocol analysis procedure
The previously unanalysed think-aloud protocols were first transcribed for each board position considered by the five novices and masters (n = 57, 3 protocols were inaudible including 2 random board positions from the grandmaster, and 1 random board position from a Fide Master). Then the protocols were broken down into segments according to guidelines indicating that each move, comment about the nature of the position, and implicit signifiers of evaluation should be put into a different segment (Newell & Simon, 1972). An example of a segmented protocol is given in Table 2 below:

Table 2: An example of a segmented expert protocol. This protocol corresponds to the fourth master problem-behaviour graph in Appendix C.

Grandmaster (participant 4), normal position 1, black
1.   So, I'm black in this position.
2.   It's some kind of Alekhine Defense,
3.   Caro-Kann something in that line,
4.   black to play
5.   and white has immediate threat f5
6.   at the same time all of my pieces seem to be ok,
7.   except maybe for the Bg6.
8.   So I now need to find a way to stop f5,

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

9.   I have semi-open file h,
10.   which could be leading somewhere,
11.   em I can stop f5 by playing Ne7
12.   but that could lead to Bb4
13.   threatening to,
14.   well obviously wanting to get the knight
15.   in or planting the bishop on d6,
16.   probably need to play c5,
17.   which I can do if I have to if…
18.   or next… (*time up*)

These segmented protocols were then worked through move by move and segment by segment alongside a chess board to map each move verbalized by a player for a board position. The number of segments for normal board positions ranged between 32 to 74 segments for masters and 27 to 48 segments for novices. The length for random board positions ranged between 30 to 64 segments for masters and 24 to 54 for novices. Problem behaviour graphs were constructed from each protocol indicating the order of moves and move sequences considered with their retrospective evaluations (see Appendix D for a subset of problem behaviour graphs generated).

Scoring

The data was scored in relation to the primary measure, the sorts of hypothesis tests, falsification or confirmation, that players carried out. It was also scored in relation to six search behaviour measures including, (i) first move quality, (ii) number of individual moves, (iii) search depth, (iv) search breadth, (v) and the different sorts of articulated move sequence: complete, skipped moves, base skip move, and ambiguous; (vi) point at which skipped moves occur.

To illustrate the scoring, two problem behaviour graphs are presented, one constructed for a grandmaster and one for a novice for position 1, when it is black to play in Figure 3 and Figure 4. A visual comparison of the grandmaster graph to the novice graph, illustrates their differences. The most obvious difference is the size. The master's graph has more moves than the novice's. These differences will be explored in more detail shortly.
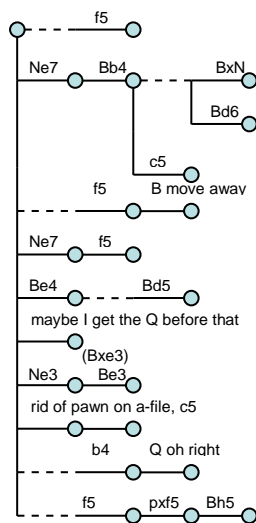
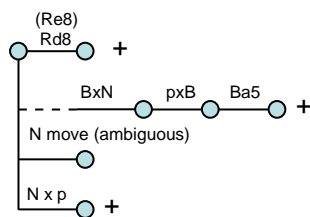Figure 3: Grandmaster (participant 4), normal position 1, black

Figure 4: Novice (participant 12), normal position 1, black

First, we will present a step-by-step interpretation for the grandmaster's graph and outline what we mean by different types of moves sequences using examples from the graph (e.g., what we mean by a complete, skipped, baseskip or ambiguous move sequences). Search is from left to right and then down. The circles represent nodes and each node indicates where the representation of the position has changed due to the imagination of a move that has taken place. The top-left-hand node represents the starting position. This problem behavior graph corresponds to the think-aloud protocol we presented in Table 2. The graph shows that the Grandmaster's first move for board position 1 (where he had the black pieces and it was black to move) is indicated as 'f5' in the graph. The 'f5' move refers to a pawn being imagined to move to the f5 square. The full verbalization was 'and white has immediate threat f5' (segment 5 in Table 2) and the pawn move is for the opponent (white). It is clear that it is an opponent move because it is a move that is only possible for white. The dashes preceding f5 on the graph indicate that a move has been skipped, that is, the Grandmaster skipped his own turn to move in his verbalization. In the graphs a dash (---) represents a move for the opponent or for the self was not articulated (a 'skipped' move). The next verbalized move is segment 11, 'em, I can stop f5 by playing Ne7'. Ne7 refers to the black knight moving to the e7 square. Then the move Bb4 is verbalized for white to follow this move: 'but that could lead to Bb4'. This verbalization is followed by skipping a move for the self and instead verbalizing two possible moves for white after Bb4. Notice the dashes on the problem behavior graph to indicate the skip. Both of the possibilities considered are

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

connected to what the grandmaster thinks white's plans are following his Bb4 move. First, 'well obviously wanting to get the knight', which refers to the white bishop that is now poised to capture his black knight on e7. This is indicated as BxN on the graph, where the 'x' is the symbol used to indicate 'capture'. Second, 'or planting the bishop on d6' refers to the other possibility for white of playing his bishop to the d6 square (Bd6).

In the graphs ambiguous move sequences are recorded using the player's verbalization, for example, 'R moves' implies the Rook moves, but we do not know where it moves to so it is labeled 'ambiguous' (for an example of an ambiguous move sequence, see master's problem behavior graph example 3, position 2, normal board position, fourth move sequence down in Appendix D). Repeated move sequences are annotated with 'repeat'. The end nodes or terminal nodes are annotated with a plus, minus or equals sign. A plus sign (+) indicates that the player provided the retrospective evaluation that the position led to a positive outcome; a minus sign (-) indicates that a player evaluated it as leading to a negative outcome, and an equals sign (=) indicates that a player evaluated it as leading to a neutral outcome.

Results and discussion
The results are reported for: the nine types of hypothesis test; the quality of the first verbalised move; the total number of moves; the search depth; the search breadth; the types of move sequences; and skipped moves along a move sequence.

Hypothesis tests in complete move sequences
Only completely articulated move sequences (sequences in which the players articulated every imagined move of their own and their opponent) can be reliably inputted to the computer program *Fritz* for objective evaluation. 50% of move sequences were articulated completely by masters and novices. The mean number of different types of hypothesis tests are presented in Table 3.

Table 3: The mean number of different types of hypothesis tests for the complete move sequences for the normal and random board positions by the masters and novices.

Nine test types	Normal		Random	
	Master	Novice	Master	Novice
+/+	0.60	0.27	0.67	0.27
-/- (falsification)	**1.07**	**0.40**	**0.83**	**0.87**
=/=	0.53	0.13	0.08	0.00
+/- (confirmation bias)	**0.53**	**0.87**	**0.42**	**0.33**
+/=	0.47	0.60	0.17	0.20
=/-	0.13	0.20	0.17	0.27
-/+	0.13	0.07	0.42	0.13
-/=	0.33	0.13	0.17	0.07
=/+	0.13	0.20	0.08	0.20
Total	3.92	2.87	3.01	2.34

Note: The falsification and confirmation bias hypothesis tests are represented in bold.

The hypothesis test type '-/-' indicates that the player, and *Fritz,* both evaluated a move

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

sequence as leading to a negative outcome as Table 1 showed. It is the strongest falsification type of the nine tests. Masters falsified more often than novices for normal positions (1.07 vs 0.40, Mann-Whitney $U_{5,5}$ = 4.00, Z = -1.844, p = .0325), and not for random positions (0.83 vs 0.87, Mann-Whitney $U_{5,5}$ = 9.00, Z = -.752, p = .226*)*. The result corroborates the prediction that masters try to falsify their hypotheses. It also suggests that when domain knowledge is eliminated masters do not falsify more than novices.

The '+/-' test type corresponds to the strongest type of confirmation bias as Table 1 showed. The player evaluates a move sequence as leading to a positive outcome and *Fritz* evaluates it objectively as leading to a negative outcome. There were no differences in frequency of confirmation bias in masters and novices for the normal positions (0.87 vs 0.53, Mann-Whitney $U_{5,5}$ = 8.50, Z = -.854, p = .197), or the random positions (0.33 vs 0.42, Mann-Whitney $U_{5,5}$ = 7.5, Z = -1.195, p = .116).

Falsification and confirmation bias were the two most frequently occurring of the nine test types, as Table 3 shows, and they account for almost half of the total number of hypothesis tests in the complete move sequences. The nine test types are collapsed into three broad test types: objective tests, positive bias tests and negative bias tests as Table 4 shows).

Table 4: The mean number of positive, negative and objective tests by the masters and novices for normal and random board positions (with standard deviations in parentheses).

	Masters	Novices
Normal board positions		
Objective	6.0 (3.00)	2.4 (1.82)
Positive	3.2 (1.10)	5.0 (2.74)
Negative	1.8 (1.64)	1.2 (0.84)
Random board positions		
Objective	3.8 (1.79)	3.4 (1.52)
Positive	1.8 (1.64)	2.4 (0.89)
Negative	1.8 (1.92)	1.4 (1.14)

Note: Objective tests include the falsification tests, and positive tests include the confirmation bias tests.

Objective tests are tests for which there was a match between the player's subjective evaluation and *Fritz's* objective evaluation (the '+/+', '-/-' and '=/='tests). The category includes the falsification test. Masters carried out more objective tests than novices for the normal positions (6.0 vs 2.4, Mann-Whitney $U_{5,5}$ = 5.00, Z = -1.596, p = .055), but not for random positions (3.8 vs 3.4, Mann-Whitney $U_{5,5}$ = 10.5, Z = -.426, p = .335). The result again shows that masters were better at objectively evaluating their move sequences than novices, but only when they could rely on their domain knowledge.

Positive bias tests were tests for which the players were more positive than *Fritz* (the '+/-', '+/=', and '=/-' tests). The category includes the confirmation bias test. There were no differences in frequency of positive bias tests between experts and novices for the normal positions (5.0 vs 3.2, Mann-Whitney $U_{5,5}$ = 7.5, Z = -1.085, p = .139), or random positions (1.8 vs 2.4, Mann-Whitney $U_{5,5}$ = 9.0, Z = -.759, p = .224*)*.

Negative bias tests were tests for which the players were more negative than Fritz (the '-

/+', '-/=', and '=/+' tests). Masters and novices carried out the same amount of negative tests for the normal positions (1.8 vs 1.2, Mann-Whitney $U_{5,5} = 10.5$, Z = -.434, p = .332), and for random positions (1.8 vs 1.4, Mann-Whitney $U_{5,5} = 12.0$, Z = -.108, p = .457).

Overall, the primary difference in hypothesis testing between masters and novices is that masters were better at objectively evaluating their moves than novices, and in particular they were better at accurately falsifying their hypotheses than novices, when they had domain knowledge to rely on.

First move quality
Masters and novices differed somewhat in the quality of the first moves they verbalized as a possible move for them to play, as evaluated by *Fritz 8*. The Fritz evaluations correspond to how much of a pawn may be won or lost if a hypothesized move is played, for example -.99 indicates the player will lose almost a full pawn, .99 indicates that the player will win almost a full pawn. In chess terms a pawn is a considerable loss, particularly if one is losing a pawn's worth every other move. Masters' first moves tended to be of a somewhat better quality than novices for normal positions, as judged by *Fritz* (-.070 vs -.440) but the difference was not significant (Mann-Whitney $U_{5,5} = 6.5$, z = -1.257, p = .104). Masters' first moves tended to be of a similar quality to novices for random positions (-6.47 vs -5.05, Mann-Whitney $U_{5,5} = 11$, z = - .313, p = -.754, two-tailed). Earlier findings (De Groot, 1965; Gobet & Campitelli, 2002) suggest that the first move is generated by masters directly from their long-term knowledge, for example by accessing a template (Gobet, 1998; Gobet et al., 2004). In our experiment the masters did not generate significantly better first moves than the novices in the normal positions. This suggests that hypothesis testing may often help masters think about the position further in order to generate better moves for play. Further searching may help them to falsify their first move and trigger better candidate moves. Hypothesis falsification may help masters to avoid playing moves that would lead to error, even though they generated the moves.

Number of moves
In the protocol analysis we can now calculate the average number of *individual* moves per position. Each move sequence contains different numbers of individual moves, for example one move sequence may contain one move whereas another may contain eight individual moves. Masters generated more individual moves than novices in normal positions (25.40 vs 16.80) although the difference is marginal (Mann-Whitney $U_{5,5} = 4.0$, Z = -1.776, p = .076, two-tailed), and they did so in random positions too (31.56 vs 18.27, Mann-Whitney $U_{5,5} = 2.0$, Z = -2.193, p = .028, two-tailed) as Table 5 shows. The result suggests that masters can generate a larger search space than novices.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Table 5: The mean ply depth and ply breadth of move sequences, and the mean number of individual moves in the generated move sequences (standard deviations in parentheses).

	Masters	Novices
Normal board positions		
Individual moves	25.40 (6.757)	16.80 (9.001)
Ply depth	3.30 (0.961)	2.61 (0.621)
Ply breadth	2.20 (1.523)	1.60 (1.230)
Random board positions		
Individual moves	31.57 (4.806)	18.20 (7.596)
Ply depth	4.13 (1.181)	2.94 (0.520)
Ply breadth	1.93 (0.864)	1.00 (0.848)

Search depth

Search depth refers to the number of moves a player thinks about in advance when choosing a move for play. These moves combined together form move sequences. The first move in the sequence is one move deep. The second move is two moves deep. Search depth is measured by the number of moves that indicates the length of a move sequence. Each move in the sequence is given a measure of one ply. For example the search depth of a move sequence that is three moves long is termed a '3 ply' move sequence. Consider problem behavior graph example 6 in Appendix D. In the third move sequence down the novice first considers playing his pawn to the f4 square (f4; 1 ply). Then he thinks about his opponent responding to this move by capturing this pawn on f4 with a pawn placed on an e-square (exf4; 2 ply). Now he thinks about another move for his opponent immediately (indicating a skipped move for himself; 3ply), and he thinks about the opponent playing the queen to the e2 square capturing one of his pawns and lining up for an attack on his king putting him in check (Qxe2+; 4 ply). This move sequence is four moves deep and is called a 4 ply move sequence. Notice that skipped moves are given one ply also in order to represent the search depth as accurately as possible. We can calculate search depth for all move sequences, whether they were complete move sequences (every move was articulated precisely along the sequence), skip move sequences (at least one move for the self or opponent was not articulated along the sequence), base skip move sequences (the first move of the sequence was not articulated) or ambiguous move sequences (the coordinates of at least one move were not articulated).

   Masters and novices searched to the same depth in normal positions (3.3 vs 2.61, Mann-Whitney $U_{5,5}$ = 7.0, Z = -1.149, p = .251, two-tailed), although masters searched to a somewhat greater depth than novices in random positions (4.14 vs 2.94, Mann-Whitney $U_{5,5}$ = 4.0, Z -1.776, p = .076, two-tailed), as Table 5 shows. Previous studies have not found a difference in search depth. For example, DeGroot (1965) did not find that masters searched more than lesser experts *(*in a study of five grandmasters, five experts, and five category A players), and the result has been replicated (e.g., De Groot & Gobet, 1996; Gobet, 1998; Gobet et al., 2004). However, a recent study has shown that masters (n = 2) search deeper in board positions that are highly complex, that is, positions that can occur in real games and require at least 23 look-ahead moves to solve (Gobet & Campitelli, 2002). The current study provides a detailed comparison of search between normal and randomized chess positions comparing a large sample (n = 10) of master and novice players, with a large skill difference between the two groups. It replicates previous findings of no search depth difference for

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

normal positions, but suggests that masters may be able to search deeper in the random positions.

Search breadth

Search breadth refers to the situation in which more than one possible move has been considered at the same place in a move sequence for either the opponent or for the self. In other words, a move sequence 'broadens out' where two or more distinct alternative move sequences follow on from the same first moves. Consider problem behaviour graph example 2 in Appendix D. In the seventh move sequence down this master first considers playing his pawn to the e3 square (e3). Then he thinks about his opponent responding to this move by playing a knight to e6 (Ne6). Now he thinks about playing two alternative moves in response to his opponent's knight move: queen to the c2 square (Qc2) or queen to the d5 square (Qd5). Both of these queen moves are the third move of the sequence. On the graph this broadening (or branching) is represented by inserting the second alternative move directly underneath the first resembling a fork shape. The number of prongs of the 'fork' represent the number of moves imagined possible at the same stage in a sequence. We can calculate search breadth for all move sequences.

Masters searched as broadly as novices for normal positions (2.2 vs 1.6, Mann-Whitney $U_{5,5} = 9.0$, $Z = -.760$, $p = .459$, two-tailed) but they searched somewhat more broadly than novices for random positions (1.93 vs 1.00, Mann-Whitney $U_{5,5} = 4.5$, $Z = -1.735$, $p = .083$, two-tailed) as Table 5 shows. The result suggests that when masters are confronted with a novel problem in their domain (random chess positions) they may be able to rely on superior search driven processes.

Types of move sequences

As reported above, 50% of move sequences were completely articulated sequences. We identified three other types of move sequence from the problem behavior graphs. As Table 6 shows, the four types of sequences are complete, skipped, base skip, and ambiguous sequences and we describe each in turn. Masters and novices produced the same amount as each other of each type of move sequence for the normal and random positions, as Table 6 shows.

Half of all move sequences were *complete move sequences* where every move for the player and his or her opponent was articulated. For example, the problem behavior graph in Appendix D, example 2 (International Master, position 1 with white to move) provides a good illustration of a complete sequence in the seventh move sequence down. The master thinks about moving his pawn to the e3 square (e3), then he thinks about his opponent playing his knight to the e6 square (Ne6), and then he thinks about responding to this opponent move by playing his queen to the c2 square (Qc2).

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Table 6: The mean number of different sorts of move sequences generated by the five masters and five novices for the normal and random board positions (standard deviations are in parenthesis).

Move sequence	Complete	Incomplete	Base skip	Ambiguous
Normal positions				
Master	4.00 (1.13)	2.20 (1.02)	1.67 (0.53)	0.33 (0.34)
Novice	3.13 (0.96)	1.53 (1.41)	1.14 (0.51)	0.60 (0.69)
Total	*3.57 (1.09)*	*1.87 (1.21)*	*1.40 (0.56)*	*0.47 (0.53)*
Random positions				
Master	2.90 (1.84)	2.76 (0.70)	1.30 (0.77)	1.37 (0.71)
Novice	2.67 (1.25)	1.60 (0.72)	1.20 (0.81)	1.53 (0.51)
Total	*2.79 (1.49)*	*2.18 (0.91)*	*1.25 (0.75)*	*1.45 (0.59)*

Note. Means are based on 15 cases (five players by three board positions) except for the masters' random positions which were based on 12 cases because of a recording error.

Then the sequence ends and is evaluated (negatively). There are no moves left out of the sequence for either the player or his opponent. Masters and novices generated the same number of complete move sequences for normal positions (4.00 vs 3.14, Mann-Whitney $U_{5,5}$ = 7.0, Z = -1.16, p = .246, two-tailed) and for random positions (2.90 vs 2.67, Mann-Whitney $U_{5,5}$ = 12.0, Z = -.105, p = .917, two-tailed). The results indicate that masters and novices were equally able to articulate their own and their opponent's moves.

The next most common sort (25%) were *skipped move sequences* where an essential move was not mentioned by the player at some point in the move sequence. These incomplete skipped move sequences had from one to four skips in the sequence. For example, as Appendix D, example 3 shows in the fifth move sequence down the master thought first about moving his queen to the e5 square (Qe5). He then thinks about moving his knight to e4 directly after his queen move and therefore skips a move for his opponent to respond to his queen move. This skip is represented as three dashes between the first and third moves. The fourth move in the sequence is where he thought about the opponent capturing his knight that has just moved to e4. He thinks about his opponent capturing his knight with a bishop (Bxe4). The final and fifth move in the sequence is his response to this capture. He thinks about capturing the opponent bishop that has just taken his knight with a pawn. The move sequence is now terminated. Masters and novices generated the same number of skipped move sequences for normal positions (2.2 vs 1.53, Mann-Whitney $U_{5,5}$ = 8.0, Z = -.946, p = .344, two-tailed), and masters generated more than novices for random positions (2.76 vs 1.60, Mann-Whitney $U_{5,5}$ = 3.0, Z = 2.015, p = .044, two-tailed). The results indicate that masters and novices sometimes have difficulty in articulating their own and their opponent's moves, but masters have this difficulty most in randomized positions were their expert knowledge is absent.

A third common sort of move sequence (19%) was *base skip sequences* where the first move or 'base move' of the sequence was not mentioned. For example, as Appendix D, example 2 shows in the first move sequence of the graph this master thought first about the opponent moving a knight to the e6 square (Ne6) even though it is the master's turn to move. He has skipped his move and it is represented as three dashes at the start of the move sequence. The opponent Ne6 move is placed as the second move in the sequence. The third

move in the sequence is where he thinks about moving his queen to the e3 square in response to his opponent's move. The fourth move in the sequence is where he thought about the opponent responding by moving his queen to the b5 square (Qb5). He thinks immediately about the next opponent move in the sequence rather than his own. He thinks about the opponent moving his rook to the e8 square (Re8). The move he skipped for himself is the fifth move in the sequence, and is represented by three dashes between the fourth and sixth moves in the sequence which both belong to the opponent. The move sequence is now terminated. Masters and novices generated the same number of base skip move sequences for normal positions (1.67 vs 1.14, Mann-Whitney $U_{5,5} = 5.5$, $Z = -1.49$, $p = .136$, two-tailed) and for random positions (1.3 vs 1.2, Mann-Whitney $U_{5,5} = 11.5$, $Z = -2.10$, $p = .834$, two-tailed). The results show that masters and novices sometimes do not articulate their first moves, and this difficulty occurs for both normal and random positions.

Finally, some move sequences were ambiguous (6%), where the move sequence could not be interpreted. For example a player said 'then the rook moves over there…', and it is not possible to determine the precise coordinate the player has in mind for it to move to. Masters and novices generated as many ambiguous sequences for normal positions (0.33 vs 0.60, Mann-Whitney $U_{5,5} = 10.5$, $Z = -.435$, $p = .663$, two-tailed) and for random positions (1.37 vs 1.53, Mann-Whitney $U_{5,5} = 11.5$, $Z = -.211$, $p = .841$, two-tailed).

Overall masters' and novices' think-aloud protocols of their move sequences reveal completely articulated sequences, skipped move sequences, base skip sequences and ambiguously articulated sequences. Masters and novices did not differ systematically in the frequency of each sort of sequence, but masters did generate more skipped moves in their sequences for randomized positions.

Hypothesis tests and skipped move sequences

For the normal board positions, players failed to verbalize a move in 25% of move sequences (N = 56); some move sequences had more than one such 'skip' and so there were 75 skips overall. For the random board positions, players failed to verbalize a move in 26% of sequences (N = 57), but some move sequences had more than one skip, and there were 161 skips in total. Masters and novices produced a similar number of *individual skips* in normal positions (3.064 vs 1.932, Mann-Whitney $U_{5,5} = 7$, $Z = -1.160$, $p = .246$, two-tailed) and in random positions (7.466 vs 4.2, Mann-Whitney $U_{5,5} = 5$, $Z = -1.581$, $p = .114$, two-tailed).

Masters and novices skipped their own moves as often as each other for normal positions (.40 vs .20, Mann-Whitney $U_{5,5} = 7.5$, $Z = -1.107$, $p = .268$, two-tailed); they also skipped their opponent moves as often as each other (2.67 vs 1.73, Mann-Whitney $U_{5,5} = 7.5$, $Z = -1.051$, $p = .293$, two-tailed). It is notable that they also skipped their own moves for random positions as often as each other (3.1 vs 2.0, Mann-Whitney $U_{5,5} = 6.5$, $Z = -1.261$, $p = .207$, two-tailed) and they skipped their opponent moves as often as each other (4.37 vs 2.20, Mann-Whitney $U_{5,5} = 5$, $Z = -1.586$, $p = .113$, two-tailed). Masters and novices tended to articulate their own moves more than their opponent's moves, which may reflect a sort of confirmation tendency.

To test whether it is a confirmation bias, the first few moves of a skipped move sequence was inputted to *Fritz* until the position was reached at which the first skipped move occurred. An analysis was conducted only for the normal positions, because for the random positions the large number of skips precluded inputting them. At this point (first skip node) we sought an objective evaluation of the position from *Fritz*. Masters and novices sometimes began to skip at a point in the move sequence when the evaluation of the node was negative (4.4 vs 2.4, Mann-Whitney $U_{5,5} = 324.5$, $Z = -1.083$, $p = .279$, two-tailed); they also sometimes

began to skip at a point in the move sequence when the evaluation of the node was positive (2.2 in each case). The *Fritz* evaluation for the first skipped move was somewhat more negative for masters than novices (-.6167 vs .1935) although the difference was not reliable (t = -1.620, df = 54, p = .111, two-tailed). The tendency to skip opponent's moves may simply reflect a richer representation of own hypotheses and plans.

Summary

The main result of the experiment is that masters objectively evaluated their moves more accurately than novices, and in particular they falsified accurately more often than novices for normal positions, and not for random positions. The result corroborates the prediction that masters try to falsify their hypotheses. It also suggests that when domain knowledge is eliminated masters do not falsify more successfully than novices.

The analysis also explored the differences in the search structure of master and novices hypothesis testing, and provided the following results:

(i)   Masters and novices did not significantly differ in the quality of their first moves.

(ii)   Masters generated reliably more individual moves than novices.

(iii)   Masters searched to the same depth as novices for normal positions, but to a marginally greater depth for random positions.

(iv)   Masters searched to the same breadth as novices for normal positions, but to a marginally greater breadth for random positions.

(v)   Masters' and novices' think-aloud protocols contained completely articulated sequences, skipped move sequences, base skip sequences and ambiguously articulated sequences. Masters and novices did not differ systematically from each other in the frequency of each sort of sequence.

(vi)   An analysis of the skipped move sequences shows that masters and novices reliably tend to articulate their own moves more than their opponent's moves. Masters and novices sometimes began to skip at a point in the move sequence when the evaluation of the node was negative and sometimes when it was positive. Even though skipped moves were as often for self and opponent play for masters and novices, masters generated significantly more move sequences with a skipped move for randomized as opposed to normal positions.

Let us turn now to a discussion of the implications of the results. First, how the results on search differences between masters and novices may contribute to our understanding of hypothesis testing and theories of reasoning are discussed. Second, the implications of the results for theories of cognitive expertise, and the development of expertise are implicated.

General Discussion

The results of the experimental analysis of chess masters hypothesis testing corroborate the prediction that masters try to falsify their hypotheses. They also indicate that when domain knowledge is eliminated masters do not falsify more successfully than novices. The results are suggestive that domain expertise may help experts to choose better quality moves, and the process by which it may help them is through falsification.

There was no difference in the quality of their first moves. Masters generated more individual moves than novices. They searched to the same depth as novices for normal positions, but to a somewhat greater depth for random positions. They also searched to the same breadth as novices for normal positions, but to a somewhat greater breadth for random positions. The results are suggestive that masters may have superior search skills than novices

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

for dealing with novel problems in their domain.

Masters and novices' did not differ systematically from each other in their ability to articulate their move sequences. The novices in this study were experienced novices, to ensure they were fluent in 'chess notation'. Half of all masters and novices' think-aloud protocols contained completely articulated sequences, and the remainder contained skipped moves, that is, moves that were not articulated, either during the sequence (skipped move sequences) or at the outset (base skip sequence). Masters and novices both tended to articulate their own moves more than their opponent's moves.

Implications for Theories of Cognitive Expertise

Is chess expertise comparable to other domains of expertise? Chess is adversarial and falsification in chess is a result of considering opponent moves. There are immediate consequences of committing an error: you will lose the game. Science too is a competitive enterprise and refutations may often come from competing scientists and laboratories (Kuhn, 1993; Mitroff, 1974; Gorman, 1995a). In fact, findings from the domain of chess have an excellent external validity (e.g., Larkin et al., 1980; Lesgold, Rubinson, Feltovich, Glaser, Klopfer, & Wang, 1988; Sloboda, 1976).

A hypothesis testing model implies that masters sometimes consider move sequences that lead to error. The idea that expert knowledge sometimes contains errors, or that the retrieval of knowledge sometimes results in errors implies that experts must be able to detect errors that could result from such inaccurate knowledge or inaccurate retrieval. Although expertise may be reproductive in routine problems (e.g., Ericsson, Krampe, & Tesch-Romer, 1993; Gobet 1998), masters still search beyond the first move considered and the purpose of this search requires explanation (Charness, 1991).

The addition of a hypothesis testing component to chess expertise provides the explanation that although masters may rely on templates stored in memory, the plans or moves stored with the template may not fit with every position encountered. A master may avoid error by searching to check the alternative moves retrieved as possible plans. The previous analyses of the search process in chess have focused on surface structures of the search tree, such as ply-depth and ply-breadth but these features may be only peripherally related to processing differences in search between masters and novices. A more important feature of the search structure may be the evaluation of move sequences, which may contribute to the selectivity of search.

The problem of selectivity in search is relevant for chess playing programs in artificial intelligence research (e.g., Hsu, Campbell & Tan, 1997). Chess playing programs still search extensively through the problem space when choosing a move for play. For example, the *Deep Blue* program that defeated the reigning world chess champion Kasparov considered 90 billion moves at each turn, at a rate of 9 billion per second (Eysenck & Keane, 2000). Computer chess programs today still do not perform to world champion standard without employing extensive search. The processes of chess playing programs at this level of expertise do not yet match the types of processes relied on by human world champion chess players (Hsu, Campbell, & Hoane, 1995).

Evaluative knowledge, including the goal to search for negative evaluation when searching through a problem space, as proposed by this hypothesis testing model, creates a highly selective search strategy in two ways. First, the ply depth is shortened by searching for opponent moves that falsify a plan as early in the move sequence as possible. This strategy eliminates redundant alternative lines where a poor quality opponent move that leads to a good outcome is considered even though an opponent move that falsifies it is available. If a poor quality opponent move is considered in a move sequence instead of a falsifying

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

opponent move the player may have to search deeper or broader before they find evidence to act as a 'stop rule' for their consideration of this particular move sequence. Masters' search is more selective because it is limited largely to finding opponent-falsifying moves. Second, masters' stop rule for search is to stop when the opponent's falsifying move leads to a negative evaluation, that is a worsening of the masters' position. For example, masters skip opponent moves when they think-aloud at the point at which the move sequence begins to lead to a negative outcome, according to *Fritz*. A crucial principle for the improvement of chess playing programs might be to simulate the development of the human player's goal hypothesis by employing a cyclic, progressive-deepening strategy based on hypothesis testing which acts as goal feedback. The program could carry out investigative tests and collect the test results in terms of whether a hypothesis is confirmed or falsified (that is, a board goal is obtainable or not).

Masters and novices both showed some evidence of confirmation bias: they sometimes thought about how their planned move sequences could lead to an advantage for their position when objectively it led to error. The type of confirmation bias exhibited was of the +/- test type, that is, the players either did not see that the opponent moves would lead to a disadvantage or they mentally minimized the negative impact of the opponent moves and found justifications for playing the moves anyway. It is important to note that there were very few of the '+/+' test types, which shows that the players were not simply searching for opponent moves that made their plans work out (if they were then the objective evaluation of the move sequence by *Fritz* at the terminal node would also often be positive). The confirmation shown by players is more likely a bias, that is, a failure to detect that the opponent move is negative for them, rather than a direct search strategy.

Masters may consider possible moves that they have not retrieved from template knowledge but that they have constructed 'on the spot' to explore new possibilities and so update their knowledge. This strategy may ensure they discover novel moves that are better than any previous moves considered best in a given position. Masters search even after they have generated a good move in normal chess. They may create new knowledge from old knowledge by using smaller units of chunked knowledge, lower down the hierarchical architecture than templates, to generate exploratory moves (Anderson, 1983). The structure of expert knowledge in a template system can be adaptive and promote the development of new knowledge structures if an evaluation mechanism can operate, that is, a hypothesis testing process mediated by evaluative knowledge. The conjecture provides the germ of an explanation for how knowledge may advance in a domain, and how masters deal with novel problems within the domain of chess.

Some Implications for Cognitive Theories of Reasoning

The experimental analysis of the chess masters' hypothesis testing shows that chess masters are capable of falsifying their hypotheses. They thought about how their opponent might refute their plan in move sequences that objectively led to error in normal chess. Falsification in this expert domain appears to be possible, useful and rational (see Popper, 1959; 1963; Kuhn, 1993; and Poletiek, 1996). Falsification in thinking requires thinking about negative instances. A negative instance may be counter to or inconsistent with a mental representation currently under consideration. People tend to have difficulty in representing negation, and as a result their thinking may display a bias (e.g., Evans, Newstead, & Byrne, 1993; Newstead et al., 1992). It may require more working memory resources to maintain it (e.g., Evans, 1989). People may have a tendency to represent possibilities that are consistent and true rather than false (Johnson-Laird & Byrne, 2002). To represent negation they may have to think about alternative possibilities as well as the negative one (Johnson-Laird & Byrne, 1991). In this

experiment, masters' ability to falsify is greater for normal positions than random ones, suggesting that falsification is mediated by domain knowledge. Their stored domain knowledge may contain some counterexamples that are typical in specific positions. Masters may actively seek counterexamples in their reasoning about the possibilities for play (see Byrne, 2005; Byrne, Espino & Santamaria, 1999; Johnson-Laird & Byrne, 1991; 2002). Indeed, the ability to falsify, and consider what may be false or negative may be part of what makes an expert, and helps them to avoid making mistakes. Novices' hypothesis testing may be similar to people's everyday reasoning when they do not have expertise in a domain. The implication is that people in general may find it easier to reason with positive information, such as true possibilities, and find it difficult to reason with negative information, such as false possibilities, and the results of this experimental analysis are consistent with this suggestion (Johnson-Laird & Byrne, 1991; 2002).

In addition, chess is an action domain (Mynatt, Doherty & Dragan, 1993). Players may generate a counter move to an anticipated counter move by a process of strategic reasoning at progressively deeper levels (e.g., Camerer, 2004; Hedden & Zhange, 2002; Zhang & Hedden, 2003; 2003). Action domains require consideration of a number of alternative paths towards solution, and there is much speculation on how the consideration of alternatives plays a role in more general facets of thinking such as counterfactual possibilities in imaginative thought (e.g., Byrne, 2005), thinking about alternative causes of an event (e.g., Goldvarg & Johnson-Laird, 2001), and thinking about how things could have worked out better when one makes an error (e.g., Roese & Olson, 1995). The results of this experimental analysis imply that it may be necessary to switch attention in the search space from one move sequence to another, in order to evaluate alternative possible plans of action. As experts have better problem representations due to chunked or template structured knowledge, it may be easier to switch between alternatives and to maintain the results of previous tests (Ericsson & Kintsch, 1995). Novices on the other hand may find it difficult to disengage from a current line of investigation and may find switching to another line of investigation difficult (e.g., Cowley, 2002; Cowley & Byrne, 2004). It may also be difficult for novices to retrieve the results from earlier tests, and the same may be true for peoples' reasoning in general. When people do not have expertise in a domain they may find it difficult to generate and represent many needed alternative possibilities, because they do not have the knowledge to generate an alternative, or because they do not have the practice built up that allows people to mentally represent more information when they have domain expertise (e.g., Baddeley, 1999; Ericsson & Kintsch, 1995).

A further important feature prominent in the problem behaviour graphs of chess thinking was where a player 'skipped' a move. A skip represents a move in the sequence that was not verbalised for either the opponent or the player. In other words these skips may represent where a move has not been explicitly represented as have other verbalised moves in the sequences. The skips appeared to occur most often for opponent moves, for both masters and novices. This may imply that it is more difficult to represent an opponent's plan or an alternative plan to one's own in an explicit way. Theories of reasoning suggest that mental representation may be explicit or not explicit, given that people may reason using mental representations (e.g., Craik, 1943; Johnson-Laird, 1983). Our detailed chess problem behaviour graphs may testify that chess players may reason by constructing mental representations that are explicit in some respects (see DeGroot, 1963; Newell & Simon, 1972).

This experimental analysis has shown that not only is the consideration of multiple alternative possible moves important to hypothesis falsification, but that it is important to represent these alternative possible moves as explicitly as possible. Chess masters falsified

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

their hypotheses, by searching for alternative moves an opponent could play and these alternative moves were explicitly represented. Expert knowledge prompts the falsification of one's own hypotheses, regardless of the availability of falsifying evidence (Klayman & Ha, 1987). Chess masters must consider an opponent, but it is not clear whether this competitive variable is affecting how they test their hypotheses apart from their access to large repositories of domain knowledge. This experimental analysis shows that expert knowledge is the main factor facilitating falsification in chess masters' hypothesis testing. People can experience falsification as a possible and rational strategy; chess masters use hypothesis falsification to avoid error in their thinking (Poletiek, 1996; 2001).

By examining the search process in chess masters' hypothesis testing, theories of cognitive expertise may be able to move beyond theoretical frameworks grounded in systematic pattern recognition (e.g., Chase & Simon, 1973a; 1973b), which relies on the assumption that cognitive expertise is largely reproductive (e.g., Gobet; 1998; Ericsson et al., 1993). The results in this paper suggest that there is a more adaptive view of the interaction between the expert and the problem at hand, for example, how novel an expert judges a problem to be. Little is known about how experts adapt their knowledge to novel problems in order to acquire new knowledge structures (Tei Leine & Saariluoma, 2001), detect errors (e.g., Green & Gilhooly, 1992), or learn from experiencing mistakes (Frensch & Sternberg 1993).

The hypothesis testing component of expert thinking aims to understand how experts detect errors, recover from them and adapt to novel problems in order to create new knowledge. Many disciplines require experts to test hypothetical predictions or solutions to problems whether they are scientists (e.g., Kuhn, 1993; Fugelsang et al., 2004), legal experts (e.g., Britton, 1997), medics (e.g., Koriat et al., 1980) or even creative individuals (e.g., Eysenck, 1995).

In this paper's conceptualization of chess players' hypothesis testing it became apparent that not only search, but the evaluation of search results in particular was involved in rational hypothesis testing. When a planned move objectively led to a negative outcome, novices tended only to evaluate how their move could be confirmed, but masters saw how their moves could be refuted. Previous research has not distinguished between the search and evaluation process in hypothesis testing. Some theorists see search and evaluation of the result as one and the same process (e.g., Poletiek, 2005; Howson & Urbach, 1993). But in this paper the search facilitates the application of expert knowledge to evaluation. While an evaluative search can lead to an objectively falsifying result for masters, the interpretation of this falsifying result can be biased for novices.

Bibliography

Allport, G. W. (1979). The nature of prejudice (2nd Ed.). London: Addison-Wesley.

Anderson, J. R. (1983). *The architecture of cognition*. Harvard: Harvard University Press.

Anzai, Y., & Simon, H. A. (1979). The theory of learning by doing. *Psychological Review, 86*, 124-180.

Aronson, E. (1999). *The Social Animal*. New York: Worth/ W. H. Freeman.

Baddeley, A. D. (1999). *Essentials of Human Memory*. Sussex, UK: Psychology Press.

Beevor, A. (1998). *Stalingrad*. London: Viking.

Britton, P. (1997). *The Jigsaw Man*. UK: Bantam Press.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: John Wiley and Sons.

Burns, B. D. (2004). The effects of speed on skilled chess performance. *Psychological*

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Science, 15, 442-447.

Byrne, R. M. J. (2005). *The Rational Imagination*. Cambridge, MA: MIT Press.

Byrne, R. M. J., Espino, O., & Santamaria, C. (1999). Counterexamples and the suppression of inferences. *Journal of Memory and Language, 40*, 347-373.

Byrne, R. M. J., & Mc Eleney, A. (2000). Counterfactual thinking about actions and failures to act. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1318-1331.

Byrne, R. M. J., & Walsh, C. A. (2005). Resolving contradictions. In V. Girotto & P. N. Johnson-Laird (Eds.), *The Shape of Reason: Essays in honour of Paolo Legrenzi*, 91-105. Sussex, UK: Psychology Press.

Byrne, R. M. J., & Tasso, A. (1994). Counterfactual reasoning: Inferences from hypothetical conditionals. In A. Ram & K. Eiselt, (Eds.), *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, 124-120. Hillsdale, NJ: Erlbaum.

Camerer, C. F. (2004). *Behavioral Game Theory: Experiments in strategic interaction*. Princeton NJ: Princeton University Press.

Carnap, R. (1950). *Logical foundations of probability*. Chicago, IL: University of Chicago Press.

Carroll, L. (1994). *Alice's adventures in wonderland*. London: Puffin books.

Chabris, C. F., & Hearst, E. S. (2003). Visualisation, pattern recognition, and forward search: Effects of playing speed and sight of the position on grandmaster chess errors. *Cognitive Science, 27*, 637-648.

Charness, N. (1991). Expertise in chess: The balance between knowledge and search. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits*, 39-63. Cambridge, England: Cambridge University Press.

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

Chase, W. G., & Simon, H. A. (1973b). The mind's eye in chess. In W. G. Chase (Ed.), *Visual Information Processing*. New York: Academic Press.

Cherubini, P., Castelvecchio, E., & Cherubini, A. M. (2005). Generation of hypotheses in Wason's 2-4-6 Task: An information theory approach. *The Quarterly Journal of Experimental Psychology,* forthcoming.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in physics problem solving. In R. J. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence(Vol. 1),* 1-75. Hillsdale, NJ: Lawrence Erlbaum.

Christensen-Szalansky, J. J. J., & Bushyhead, J B. (1981). Physicians' use of probabilistic information in a real clinical setting. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 928-935.

Craik, K. (1943). *The nature of explanation*. Cambridge, England: Cambridge University Press.

De Groot, A. D. (1965). *Thought and Choice in Chess*. The Hague: Mouton.

De Groot, A. D. (1969). *Methodology: Foundations of inference and research in the behavioural sciences*. The Hague: Mouton.

Dunbar, K. (2000). What scientific thinking reveals about the nature of cognition. In Crowley, K., Schunn, C. D., & Okada, T. (Eds.), *Designing for Science: Implications from Everyday, Classroom, and Professional Settings.* LEA. Hillsdale: NJ.

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review, 85*, 386-416.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Elio, R., & Pelletier, F. J. (1997). Belief change as propositional update. *Cognitive Science, 21*, 419-460.

Elo, A. (1978). The rating of chess players, past and present. New York: Arco.

Ericsson, K. A., & Kintsch, W. (1995). Long-Term Working Memory. *Psychological Review, 102*, 211-245.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Ericsson, K. A., Krampe, R. Th., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363-406.

Evans, J. St. B. T. (1989). *Bias in Reasoning*. Hove, UK: Erlbaum.

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human Reasoning: The Psychology of Deduction*. Hove, UK: Lawrence Erlbaum.

Eysenck, H. J. (1995). *Genius*. Cambridge, UK: Cambridge University Press.

Eysenck, M. W., & Keane, M. T. (2000). *Cognitive Psychology: A student's handbook*. Hove, UK: Psychology Press.

Farris, H., & Revlin, R. (1989). The discovery process: A counterfactual strategy. *Social Studies of Science, 19*, 497-513.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Frank, A. (2001). *The diary of a young girl*. London: Penguin books.

Franklin, S. P. (1998). *Artificial Minds*. Cambridge, MA: MIT Press.

Frensch, P. A., & Sternberg, R. J. (1991). Skill related differences in game playing. In R. J. Sternberg (Ed.), *Complex Problem Solving: Principles and mechanisms*. USA: Lawrence Erlbaum Associates.

Fugelsang, J., Stein, C., Green, A., & Dunbar, K. (2004). Theory and data interactions of the scientific mind: Evidence from the molecular and the cognitive laboratory. *Canadian Journal of Experimental Psychology, 58*, 1392-1411.

Gärdenfors, P. (1988). *Knowledge in flux*. Cambridge, MA: MIT Press.

Gale, M., & Ball, L. J. (2003). Facilitation of rule discovery in Wason's 2-4-6 task: The role of negative triples. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society*, 438-443. Boston, MA: Cognitive Science Society.

Gale, M., & Ball, L. J. (2005). Dual-goal facilitation in Wason's 2-4-6 task: What mediates successful rule discovery. *The Quarterly Journal of Experimental Psychology, 59*, 1-13.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic books.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (1998). *Cognitive neuroscience: The biology of the mind*. London: W. W. Norton.

Girotto, V., Legrenzi, P., Rizzo, A. (1991). Event controllability in counterfactual thinking. *Acta Psychologica, 78*, 111-133.

Gobet, F. (1998). Expert memory: A comparison of four theories. *Cognition, 66*, 115-152.

Gobet, F., & Campitelli, G. (2002). *Education and chess: A critical review.* Unpublished manuscript.

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

Gobet, F., & Simon, H. A. (1996b). The roles of recognition processes and look ahead search in time-constrained expert problem solving. Evidence from grandmaster chess. *Psychological Science, 7*, 52-53.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Gobet, F., & Simon, H. A. (1996c). Recall of rapidly presented chess board positions is a function of skill. *Psychonomic Bulletin and Review, 3*, 159-163.

Gobet, F., de Voogt, A., & Retschitzki, J. (2004). *Moves in mind: The psychology of board games.* Hove, UK: Psychology Press.

Goldvarg, Y., & Johnson-Laird, P. N. (2001). Naïve causality: A mental model theory of causal meaning and reasoning. *Cognitive Science, 25*, 565-610.

Gooch, S. (1981). *The Secret Life of Humans*. London, UK: J. M. Dent & Sons Ltd.

Gorman, M. E. (1995a). Confirmation, disconfirmation and invention: The case of Alexander Graham Bell and the telephone. *Thinking and Reasoning, 1*, 31-53.

Gorman, M. E. (1995b). Hypothesis Testing. In S. E. Newstead & J. St. B. T. Evans (Eds.), *Perspectives on thinking and reasoning: Essays in honour of Peter Wason.* Hove, UK: Laurence Erlbaum Associates Ltd.

Gorman, M. E. & Gorman, M. E. (1984). A comparison of disconfirmatory, confirmatory and a control strategy on Wason's 2-4-6 task. *Quarterly Journal of Experimental Psychology, 36A*, 629-648.

Gorman, M. E., Gorman, M. E., Latta, R. M., & Cunningham, G. (1984). How disconfirmatory, confirmatory, and combined strategies affect group problem solving. *British Journal of Psychology, 75*, 65-79.

Green, A. J. K., & Gilhooly, K. J. (1992). Empirical advances in expertise research. In M. T. Keane & K. J. Gilhooly (Eds.), *Lines of Thinking: Volume 2*. Chicester: Wiley.

Gross, R. (1999). *Key studies in psychology*. London: Hoddon & Stoughton.

Harman, G. (1986). *Change in View*. Cambridge, MA: MIT Press.

Hartston, W., & Wason, P. C. (1983). *The Psychology of Chess*. London: Batsford.

Hedden, T., & Zhang, J. (2002). What do you think I think you think? Strategic reasoning in matrix games. *Cognition, 85*, 1-36.

Holding, D. H., & Reynolds, R. I. (1982). Recall or evaluation of chess positions as determinants of chess skill. *Memory and Cognition, 10(3)*, 237-242.

Holding, D. H. (1985). *The Psychology of Chess Skill.* New Jersey: Erlbaum, Hillsdale.

Hollander, M., & Wolfe, D.A. (1999). *Nonparametric statistical methods*. New York: John Wiley & Sons.

Howson, C., & Urbach, P. (1993). *Scientific reasoning* (2nd Ed.). Chicago: Open Court.

Hsu, F., Campbell, M. S., & Hoane, A. J. (1995). Deep Blue system overview. *In Proceedings of The Ninth International Conference on Supercomputing, 240-244*. USA: ACM Press.

Johnson-Laird, P. N. (1983). *Mental Models.* Cambridge: Cambridge University Press.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hove, UK: Lawrence Erlbaum Associates.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review, 109*, 646-678.

Johnson-Laird, P. N., Girotto, V., & Legrenzi, P. (2004). Reasoning from Inconsistency to Consistency. *Psychological Review, 111, 3*, 1-23.

Kareev, Y., & Halberstadt, N. (1993). Evaluating negative tests and refutations in a rule discovery task. *Quarterly Journal of Experimental Psychology, 46A*, 715-727.

Kerlinger, F. N. (2000). *Foundations of behavioural research* (3rd Ed.). London: Harcourt College Publishers.

King, D. (1997). Kasparov v Deeper Blue: The ultimate man v machine challenge. London: Batsford.

Klahr, D., & Dunbar, K. (1988). Dual s search during scientific reasoning. *Cognitive Science, 12*, 1-55.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Klayman, J., & Ha, Y-W. (1987). Confirmation, Disconfirmation, and Information in Hypothesis Testing. *Psychological Review, 94, 2*, 211-228.

Klayman, J., & Ha, Y-W. (1989). Hypothesis Testing in Rule Discovery: Strategy, Structure and Content. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15 (4)*, 596-604.

Klein, S. W., Wolf, S., Militello, L., & Zsambok, C. (1995). Characteristics of skilled option generation in chess. *Organisational Behaviour and Human Decision Processes, 62*, 63-69.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 107-118.

Koslowski, B. (1996). *Theory and Evidence: The development of scientific reasoning*. Cambridge, MA: MIT Press.

Kotov, A. (1971). *Think like a grandmaster.* London: Batsford.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard: Evidence from the Tower of Hanoi. *Cognitive Psychology, 22*, 143-183.

Kruglanski, A. W., & Webster, D. M. (2000). Motivated Closing of the Mind: "Seizing" and "Freezing". In E. T. Higgins & A. W. Kruglanski (Eds.), *Motivational Science: Social and personality perspectives*, 354-375. USA: Taylor & Francis.

Kuhn, T. S. (1993). *The Structure of Scientific Revolutions.* (3rd Ed.). Chicago: Chicago University Press.

Kunda, Z. (1987). Motivation and inference: Self-serving generation and evaluation of evidence. *Journal of Personality and Social Psychology, 53,* 636-647.

Kunda, Z. (2000). The Case for Motivated Reasoning. In E. T. Higgins & A. W. Kruglanski (Eds.), *Motivational Science: Social and personality perspectives*, 313-335. USA: Taylor & Francis.

Laing, R. D. (1999). *The Politics of the Family*. London: Routledge

Lakatos, I. (1970). Falsification and methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of scientific knowledge*, 91-196. New York: Cambridge University Press.

Larkin, J. H., Mc Dermott, J., Simon, D., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science, 208*, 1335-1342.

Lesgold, A. M., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise is a complex skill: Diagnosing X-ray pictures. In M. T. H. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Luria, A. R. (1987). *The Mind of a Mnemonist* (2nd Ed.). Cambridge, MA: Harvard University Press.

Mallie, E. (2001) *Endgame in Ireland*. London: Hoddon & Stoughton.

Mandel, D. R., & Lehman, D. R. (1996). Counterfactual thinking and ascriptions of cause and preventability. *Journal of Personality and Social Psychology, 70*, 450-463.

Manktelow, K. I. (1999). *Reasoning and Thinking*. Hove, UK: Psychology Press.

McKeithen, K. B., Reitman, J. S., Rueter, H. H., & Hirtle, C. (1981). Knowledge organisation and skill differences in computer programmers. *Cognitive Psychology, 13*, 307-325.

Milgram, S. (1963/1974). *Obedience to Authority*. New York: Harper Torchbooks.

Miller, G. E.(1956).The magical number 7 plus or minus 2: Some limits on our capacity for processing. *Psychological Review, 63*, 81-97.

Mitroff, I, (1974). *The subjective side of science.* Amsterdam: Elsevier.

Molloy, A. M., & Scott, J. M. (2001). Folates and prevention of disease. *Public Health Nutrition (Review), 4(2b),* 601-609.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Mynatt, C. R., Doherty, M. E., & Dragon, W. (1993). Information relevance, working memory, and the consideration of alternatives. *The Quarterly Journal of Experimental Psychology,* 46A, 759-778.

Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1977). Confirmation bias in a simulated research environment: An experimental study of scientific inference. *Quarterly Journal of Experimental Psychology, 29*, 85-95.

Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1978). Consequences of confirmation and disconfirmation in a simulated research environment. *Quarterly Journal of Experimental Psychology, 30*, 395-406.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs NJ: Prentice-Hall.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Nixon, M. (2002). *The Little Oxford Thesaurus*. Oxford: Oxford University Press.

Nunn, J. (1999). *Nunn's Chess Openings*. London: Everyman.

Oaksford, M., & Chater, N. (1994). Another look at eliminative behaviour in a conceptual task. *European Journal of Cognitive Psychology, 6*, 149-169..

Peirce, C. S. (1992). *The essential Peirce, vol. 1*. In N. Houser, C. Kloesel, & the Peirce Edition Project. Bloomington, IN: Indiana University Press.

Platt, J. R. (1964). Strong inference. *Science, 146*, 347-353.

Poletiek, F. H. (1996). Paradoxes of Falsification. *The Quarterly Journal of Experimental Psychology, 49*, 447-462.

Poletiek, F. H. (2001). *Hypothesis Testing Behaviour*. UK: Psychology Press.

Poletiek, F. (2005). The proof of the pudding is in the eating: Translating Popper's philosophy into a model for testing behaviour. In K. I. Manktelow (Ed.). *Historical and Theoretical Perspectives on Reasoning*, forthcoming.

Popper, K. R. (1959). *The Logic of Scientific Discovery*. London: Hutchinson.

Popper, K. R. (1963/1978). *Conjectures and Refutations* (4th ed.). London: Routledge and Kegan Paul.

Popper, K. R. (1992). *Unended quest: An intellectual autobiography*. London: Routledge.

Roese, N. J., & Olson, J. M. (1995). *What might have been: The social psychology of counterfactual thinking*. Hillsdale, NJ: Lawrence Erlbaum.

Saariluoma, P. (1995). *Chess Players' Thinking.* UK: Psychology Press.

Saariluoma, P., & Laine, T. (2001). Novice construction of chess memory. *Scandinavian Journal of Psychology, 42*, 137-146.

Simon, H. A., & Gilmartin, K. (1973). A simulation of memory for chess positions. *Cognitive Psychology, 5*, 29-46.

Simon, H. A., & Hayes, J. R. (1976). The understanding process: Problem isomorphs. *Cognitive Psychology, 8*, 165-190.

Sloboda, J. A. (1976). Visual perception of musical notation: Registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology, 28*, 1-16.

Snyder, M., & Swann, W. B., Jr. (1978). Hypothesis-testing in social interaction. *Journal of Personality and Social Psychology, 36*, 1202-1212.

Tukey, D. D. (1986). A philosophical and empirical analysis of subject's modes of inquiry in Wason's 2-4-6 task. *Quarterly Journal of Experimental Psychology, 38*, 5-33.

Tversky, A., & Kahneman, D. (1982). Judgment under uncertainty: Heuristics and biases. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*, 201-208. New York: Cambridge University Press.
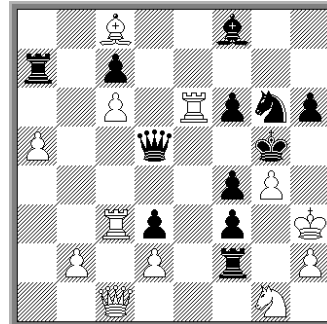
Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Tweney, R. D. (1989). Fields of enterprise: On Michael Faraday's thought. In D. Wallace & H. Gruber (Eds.), *Creative people at work: Twelve cognitive case studies,* 91-106. Oxford: Oxford University Press.

Tweney, R. D., Doherty, M. E., & Mynatt, C. R. (1981). *On Scientific Thinking*. New York: Columbia University Press.

Tweney, R. D., Doherty, M. E., Worner, W. J., Pliske, D. B., Mynatt, C. R., Gross, K. A. & Arkkelin, D. L. (1980). Strategies of rule discovery on an inference task. *Quarterly Journal of Experimental Psychology, 32*, 109-123.

Vallee-Tourangeau, F., Austin, N. G., & Rankin, S. (1995). Inducing a rule in Wason's 2-4-6 task: A test of the information-quantity and goal complementarity hypotheses. *Quarterly Journal of Experimental Psychology, 48A*, 895-914.

Van der Henst, J. B., Rossi, S., & Schroyens, W. (2002). When participants are not misled they are not so bad after all: A pragmatic analysis of a rule discovery task. *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*, 902-907. Mahwah, NJ: Erlbaum.

Van Someren, M. B., & Sandberg, J. (1994). *The Think-aloud Method*. London: Academic Press.

Vygotsky, L. S. (1986). *Thought and language* (A. Kozulin, Ed.). Cambridge, MA: MIT Press.

Waganaar, W. A., van Koppen, P. J., & Crombag, H. F. (1993). *Anchored narratives: The psychology of criminal evidence*. London: Harvester Wheatsheaf.

Walsh, C. A. (2001). *The role of context in counterfactual thinking*. Unpublished PhD Thesis. School of Psychology, Trinity College, University of Dublin.

Walsh, C. A., & Johnson-Laird, P. N. (2005). Changing your mind. *In submission.*

Wason, P. C. (1960). On the failure to eliminate hypothesis in a conceptual task. *Quarterly Journal of Experimental Psychology, 12*, 129-140.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20*, 273-281.

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. London: Batsford.

Wertsch, J. V. (1985). *Vygotsky and the social formation of mind*. Cambridge, MA: Harvard University Press.

Wetherick, N. E. (1962). Eliminative and enumerative behaviour in a conceptual task. *Quarterly Journal of Experimental Psychology, 14*, 246-249.

Wharton, C. M., Cheng, P. W., & Wickens, T. D. (1993). Hypothesis-testing strategies: Why two goals are better than one. *Quarterly Journal of Experimental Psychology, 46A,* 743-758.

Zhang, J., & Hedden, T. (2003). Two paradigms for depth of strategic reasoning in games: Response to Colman. *Trends in Cognitive Sciences*, 7, 4-5.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

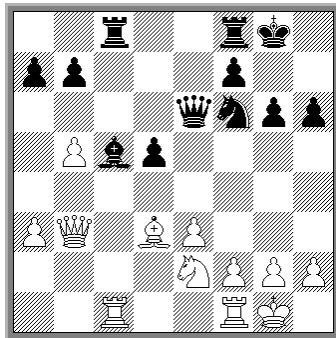**Appendix A: The three normal board positions and three random board positions used in the experiment.**
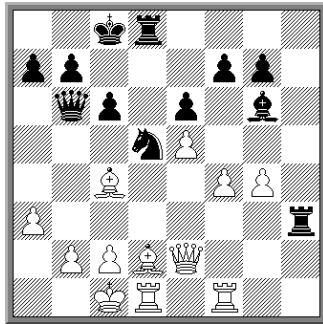
Position 1
(normal)

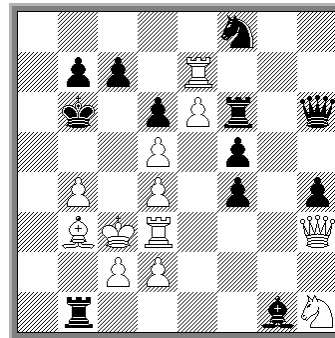**White to play** (Rook on h3 is on h8 when black to play)

Position 4
(random)

White to play (black pieces were transposed on white piece coordinates when black to play)

Position 2
(Normal)

White to play (Pawn on h6 is on h7 when black to play)

Position 5
(Random)

White to play (black pieces were transposed on white piece coordinates when black to play)
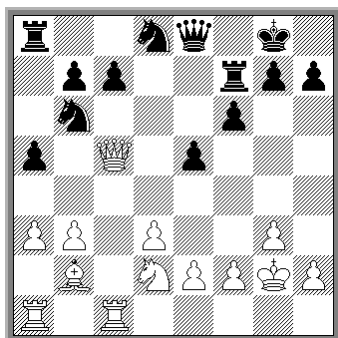
Position 3
(Normal)

White to play (Knight on d8 is on c6 when black to play)
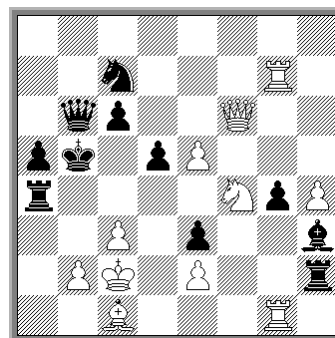
Position 6
(Random)

White to play (black pieces were transposed on white piece coordinates when black to play)

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Appendix B: Set up for chess program *Fritz 8*

The *Chessbase* analysis engine *Fritz 8* was used to estimate objective evaluations of chosen move sequences. Each node in the sequence was evaluated. These evaluations were essential in comparing chess players' expected outcomes with realistic outcome expectancies. The program has the playing strength of at least a world champion candidate grandmaster. The following procedures were followed for the computerised evaluation:

a) Maximising speed of evaluation
1) The hash table size was increased to a power of two, and was set at 256 MB to speed up the processor
2) All multimedia functions such as 'talking to you while you play' were switched off so as not to slow down the processor

b) Maximising accuracy of evaluation and reducing horizon effects
1) The openings book database was set at optimum strength
2) The number of lines of play considered in parallel was increased from two to three lines
3) Analysis was a function of ensuring at least a ply depth of 11 for each examined line using the engine's *'infinite analysis'* mode. This analysis was used to evaluate in place of a ply depth setting alone so as to counter any horizon effects likely to occur after 11ply (This analysis was undertaken by recommendation of a correspondence with professional chess grandmaster, Elo 2578. Although Chabris & Hearst, 2003 set their analysis to 10ply, and this is still very precise). The infinite analysis mode reduces horizon effects as it explores what it identifies as critical lines to a greater ply depth than others. Once the numerical evaluation stops fluctuating during the infinite analysis it is a sign that a precise estimate of the true evaluation has been reached. Estimated evaluations are expressed in 1/100ths of a pawn. The evaluation had to have at least stopped fluctuating for 15secs when the engine was processing positions at a speed of at least 750 k/Ns.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Appendix C: Segmented protocol transcripts for complete set of think-aloud data

Masters: S1, S3, S4, S8, S17
Novices: S2, S11, S12, S13, S20

Segmented protocol transcripts for normal positions

S3 Position 1: W

1. Em, the position looks so much better for white,
2. em maybe because of his sadv on the k-side,
3. em aggressive moves like f5
4. are looking quite promising.
5. Eh apart from f5 I don't seem to have that many ideas,
6. eh, I can't see anything really against f5 at the moment.
7. No, f5 seems to be critical,
8.   em exf5
9. gxf5
10. Bh5
11. em Qg2,
12. em Bxd1
13. Qxh3,
14. I've got a lot of attractive options.
15. His B has massive trouble getting out,
16. em in fact I don't think that I can,
17. all the squares seem to be covered.
18. So after f5
19. exf5
20. gxf5
21. Qg2
22. more critical for him is a move like h4
23. attacking Bc4
24. and keeping his b attacked on d1.
25. Bg5 would skewer his two rooks
26. but em to take on c4
27. doesn't seem to be too bad.
28. Em, ok I'm looking at Rh1 now
29. because I want to,
30. if I can get in f5 without any of these tricks
31. then I should definitely be clearly better in this position.
32. Em I think Rh5
33. Rgh8
34. ah doesn't seem to be a massive thing for me.
35. I could probably play Qg2
36. Rh1
37. Rdh8
38. Qg8
39. em when I get the h-file
40. possibilities of penetrating
41. doesn't seem to be too much he can do about that
42. if he Rh1
43. Rdh8
44. Qg2
45. Rxh1
46. Qxh1
47. then Qd8
48. to prevent Qa8.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

49. Its not quite clear what I've accomplished,
50. I can't play f5 in for the moment.
51. So, are these the right ideas?
52. Ne3 looks like one of his ideas if I don't do anything,
53. em forcing me to take on e3 em,
54. that position doesn't look too clear,
55. doesn't look too good for me anyway.
56. Em so I need to…

S17 Position 1: W

1. Ok, em, this is a position I would have regularly played as black,
2. em and I don't really know why I used to.
3. Em, basically white has the advantage of extra space
4. and I would regard the two bishops as being a big advantage.
5. Em, and white's only problem really is that the pawn structure is a wee bit stuck,
6. and that … em black has well controlled place on the h-file.
7. The ideas I'd be looking for… at play for white here.
8. Actually there's a few initial candidate moves.
9. I would consider playing Qe1
10. with the idea of blocking Bh5.
11. I'd … look at trying to make the tactic of playing f5…
12. doesn't work right now
13. as after exf5
14. gxf5
15. Bh5.
16. Oh…that actually might work with the …
17. probably play Qg2.
18. If white achieves that then white has again fluid pawn movement,
19. which would be good for the two bishops.
20. Eh… and an-another kind of normal idea in this position is just to eh,
21. playing in c4
22. which removes that N
23. which is well placed.
24. Em, and that would have the advantage of opening the d-file.
25. And black would be a bit vulnerable on the d-file
26. as his rook is already committed to playing on the open h-file…
27. Now the one idea that looks as though it would cut his bishop off is f5…
28. So f5
29. exf5
30. gxf5
31. Bh5
32. Qg2
33. R…
34. N black can play,
35. hmm, that's interesting…

S8 Position 1: W

1. Em, ok lets have a look.
2. Looks like Caro-Kann position.
3. Em yea, lets check the material first,
4. ya ok equal,
5. black has got the h-file.
6. W might be able to break through with f5
7. although there's a tactic…
8. aha not really

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

9. f5
10. exf5
11. gxf5
12. Bh5
13. Qg2
14. so Bxd1
15. Qxh3
16. and that looks very good.
17. Any other ideas,
18. Qg2 first, ….
19. Anything here for black
20. Rxa3,
21. no not really a threat,
22. c2 is covered,
23. Nc3
24. no I don't think I believe in all this kind of stuff.
25. Ok Qg2
26. or f5 seem to be the moves.
27. Hmm, could also think of taking on d5
28. Bxd5
29. cx
30. no that's counter play.
31. Em yea, f5
32. is there any threat at all?
33. I can't see one there
34. ok Ne3
35. Qx
36. Qx
37. f5…

S4 Position 1: B

1. So, I'm black in this position.
2. Its some kind of Alekhine Defense,
3. Caro-Kann something in that line,
4. black to play
5. and white has immediate threat f5
6. at the same time all of my pieces seem to be ok,
7. except maybe for the Bg6.
8. So I now need to find a way to stop f5,
9. I have semi-open file h,
10. which could be leading somewhere,
11. em I can stop f5 by playing Ne7
12. but that could lead to Bb4
13. threatening to,
14. well obviously wanting to get the N in
15. or planting the B on d6.,
16. probably need to play c5,
17. which I can do if I have to if…
18. or next…
19. now I think this is …
20. I don't see any other active moves at the moment…
21. nothing which comes to mind really.
22. f5 is a big threat,
23. it would be very unpleasant
24. and get the bishop to move away.
25. I guess my … Nd5…
26. ah ok, so Ne7,

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

27. what comes next?
28. So after Ne7
29. f5
30. I don't have to worry about it…
31. ideally I would like to see the bishop moving to e4
32. …d5
33. maybe I get the queen before that .
34. now, there is a problem with Ne3
35. which is Be3
36. probably because then I need now move out and get some…
37. rid of p on a-file.
38. c5 will then be much risky
39. so maybe b4 is possible
40. Q… oh right…
41. anything else here,
42. em not really
43. ok f5 for the moment is not a big threat
44. because I can take on f5
45. pxf5
46. then Bh5
47. win exchange
48. so not a big deal.
49. Now let me think,
50. got one move when I can do something.
51. What I want to do is not really clear to me…

S1 Position 1: W

1. First the position is kind of a caro-kann em type of position…
2. I would eh, like to ,
3. the knight in the former caro-Kann Bf5 variation…
4. leaving the B on g6
5. em which is obvious place …
6. but well outside of the pawn chain
7. which is supporting condensing the black pawn configuration of pawns on the light squares at e6, c6.
8. It is important to discern…that the black position …em… eh is based around two ideas.
9. They may involve Nd5
10. as the strength of it as a piece
11. which has very large em…control of the board
12. controlling indeed the whole centre…
13. and eh the black bishop on g6
14. which eh is looking towards the white k on c2
15. em it seems to be a more or less evenly balanced position…
16. its difficult to see how black is actually going to make any progress here
17. when white probably has more chances
18. due to the pair of bishops
19. em…the centrally placed rooks
20. and the threat of f5
21. which could prove annoying.
22. If black is forced to go Bh7 there …
23. then eh the bishop is very much out of play…
24. and eh then white possibly move his own Bd3
25. and perhaps with preparation play c4
26. kicking the N back
27. always gaining space…
28. em and then once the N is kicked back
29. the Be4
30. pc4

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

31. somehow manage to play his Bb4…
32. em attacking the dark squares
33. or weak squares…
34. d6…
35. and if black supports the play of b5
36. then eh c4
37. might leave the rook more exposed
38. eh…firstly white eh… think white has to be a bit careful about playing f5 too early
39. as particular f5
40. px
41. px
42. Bh5 em
43. which pins the eh… which pins the eh the queen or the em queen and rook
44. although it might be even possible now straight off is to play px
45. px
46. Bh5
47. Qg2…
48. and …blacks rook has become somewhat embarrassed
49. with QxR the black pawn…
50. or the black bishop is takes on d1
51. has no where to go
52. and white also has the idea of b6
53. or f6+ f6 discovered check
54. so win in the house.
55. f5 seems to be the critical move…
56. objectively white has possibly a set of chances here.
57. Black has got a superficially good position.
58. An active Rh3
59. but I don't see how he can make any progress against the white king
60. due to the fact that the queenside is very well consolidated by the two bishops…
61. its eh more than likely.
62. The course of play…

S3 Position 2: W
1. Ok em, isolated d pawn position,
2. equal material,
3. opposite coloured bishops,
4. em basically black's got reasonably active pieces
5. but white has definitely got the better position due to the structural weakness,
6. Nf4 attacks Queen
7. and pawn on d5,
8. em seems to lead to quite a favourable position for white.
9. Nf4
10. Qd6
11. Rfd1 seems to be logical
12. oh but the pa3 hanging in that line,
13. em so I don't really want to analyse that
14. because letting him get my a3 pawn for d5
15. which is a big let off for him.
16. Em, so do I have any other moves apart from Nf4 in this position?
17. Em Nd4 is possible
18. but not very good.
19. He can take it for one Bxd4
20. after which the structural defect is nullified
21. because I have an isolated p too.
22. But that position might be slightly in my favour as well
23. because my B is strong
24. but he doesn't have to take it.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

25. So eh I'm looking at a4,
26. a4 is prophylactic but,
27. it doesn't do much else otherwise.
28. Em, maybe maybe if I played Rfd1 immediately,
29. then if Qd6
30. just a4 there…
31. Ba3
32. eh so is a little bit annoying
33. but Rxc8
34. Rxc8,
35. yea getting in Q as well.
36. I definitely feel like I should play a move
37. just to cement my advantage here.
38. Em its not quite clear what it is.
39. His pieces are all defended em,
40. Rfd1
41. Qd6 hmm,
42. maybe Rc3
43. in that position black is better,
44. clearly mobilised his Ne4 there.
45. Rfd1
46. Qd6,
47. yea my king,
48. I'm beginning to like my position a bit less than I did at the outset.
49. Lets think
50. its quite difficult to cement an advantage here…

S17 Position 2: B

1. [Coughs]. Ok, well we have an isolated queen's pawn position (IQP)…
2. which normally means that black should be looking to try em…
3. black has the isolated queen pawn
4. and should be trying to get some activity
5. and normally in these positions you don't like to have swapped off too many pieces,
6. …and there are a couple of minor pieces missing
7. which means that I'm not happy.
8. Which means that em I'm looking perhaps at trying to equalise in this position.
9. Em, the advantages for trying to equalise includes the fact that the bishops are of opposite colours
10. and I think that em white has probably erred
11. em…by pushing the queenside pawns too far.
12. Under those kind of circumstances…em,
13. the moves that I would like to play would be…
14. initially candidate moves would be Bb6,
15. Qe5,
16. and perhaps just Qd7.
17. The ideas I would like to pursue in this position would be to gain control over the d4 square
18. so I could…pushing the pawn.
19. The other idea
20. which Qe5 kind of helps…
21. is that it allows Ne4
22. which is a good place for the N to go to.
23. It's not really in white's favour to exchange
24. so Ne4
25. Bxe4
26. pxe4 in that position
27. 'cause then em…I've got rid of my IQP
28. and em…I, I've got rid of minor pieces
29. and got a bit more space…

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

30. so actually Qe5
31. looks like a fairly decent move…
32. what kind of things can white do against that?
33. White would probably , would like to play maybe Qc3…
34. and so I would…try bring in Bb6 first
35. …white…
36. oh yea the a3 pawn is under attack
37. with Qd6 as well.
38. Qd6
39. forces a4
40. and then I can just play Na4
41. or …Qd6
42. a4
43. I can play d4…d4.
44. there is a lot to think about here.
45. Qd6
46. a4
47. Ng4
48. is there compensation developing there?
49. There's actually some very decent play developing…
50. ah now I'm beginning to see some nice ideas for…for black
51. …Nd4 just of its self,
52. em threatening taking on e3.
53. Threatening Qe5..
54. Ok…Qe5
55. Qc3
56. then that's the end of the line really,
57. of no actually I can't…

S8 Position 2: W

1. Gosh, again a middle game position.
2. I wish we would have an endgame.
3. Ok lets have a look here
4. 2-4-6, 2-4-6 ok, yea material is equal,
5. hmm b5 looks already very committal,
6. ok so isolated pawn
7. but it doesn't look too bad.
8. Yea there are some tactics in the position.
9. In principle I'd say black looks quite healthy,
10. em unless the tactic works,
11. Nf4 Nf4 hmmm,
12. I don't really like that move
13. I think its better to kind of probably play something like Bb1
14. with the idea of doubling on the d-file
15. and putting Nd4
16. and then later on try to attack pd5.
17. what I don't like about white's position is that b5 has already been made
18. and Bc5 looks very strong.
19. Does d4 work?
20. No it doesn't,
21. ok that's still pinned.
22. e4 doesn't' work.
23. Nf4
24. just attacks the Q
25. doesn't do anything.
26. Yea I think Bb1 looks like the move I would play in this position.
27. Rd1 first possible

28. but might be more,
29. might safe to protect pe3 with the Q first.
30. Hmmm after h4
31. no that's too slow.
32. Any attacks on g6?
33. No I don't see any
34. Bb1 is my move.

S1 Position 2: W

1. Eh here we have a position which clearly appears to be better for white at first glance.
2. Opposite coloured bishops
3. eh, white's only real weakness is the pe3
4. which is also in the centre…
5. the strength is its connected to the rest of the pawn chain.
6. Black e…black's isolated pd5 seems quite weak,
7. the Qe6 would not be a good defender of the pawn
8. as its exposed itself.
9. So, an immediate Nf4
10. would expose the eh…drawback of the position.
11. I suppose black does have the…
12. two ideas in this position are central.
13. First of all the idea would be to play d4 at some point
14. and exchange off the weak d pawn
15. at which point the position would be eh…would be equal.
16. For black anyway,
17. get rid of the central…
18. even though quite weak on g6
19. and h6 on the kingside
20. threatening things like…
21. there's a tactic tricks have which eh…
22. the idea of Bxe3
23. px
24. Qx+
25. Kh1
26. and eh Ng4 at that point
27. threatening ideas like eh N…
28. Rxc1
29. and Nf2+
30. could be quite embarrassing
31. although white could always throw the B in the way.
32. Nf4
33. take g pawn
34. Nh3
35. and a more or less finishing in a perpetual in some particular lines
36. on the N
37. if the N is guided certainly to g1…
38. As for what white should actually do in this position…
39. well bishops influence on b1 is not so…
40. I think if N could actually become quite a reasonable…(unable to translate for 2 seconds)…
41. not on e2 as it stands.
42. I'd tend to almost to…cover the threat of eh…to cover the threat of e3.
43. I'm reluctant to possibly play Nf4 immediately
44. as the Q goes back e5
45. which is quite a good placement.
46. The ideas Bd6
47. and e5
48. would give black some activity.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

49. But the weakness of the d pawn is not so clear then.
50. Essentially it's a structural weakness
51. black offers eh some kind of dynamic activity…
52. Perhaps doubling on the c file
53. would e eh…a way to try to…adjust the situation…
54. something tells me that if white could arrange the position to have eh h3
55. in with d4
56. could be answered with e4
57. and f4
58. and exchange of queens possibly
59. and also after d4
60. px
61. or B sac on g6
62. would no longer work then
63. as Ng4 would no longer follow.
64. So h3 might be a useful prophylactic move.
65. White actually should aim to strengthen the position gradually
66. as opposed to doing anything drastic.
67. As it is a kind of established position
68. given the play Bb1
69. em with the idea to play Rfd1
70. and Nfxd4…d4 pawn.
71. It seems a bit passive however…
72. em B would actually,
73. I think, I think h3 would be a useful move…
74. so decide…the idea…

S4 Position 2: B

1. Ok, I'm black here.
2. I've got the isolated d pawn
3. with most of my pieces rather active…
4. b pawn has gone a bit too far up tob5
5. which gives me a nice place on c5 for the B.
6. Also it …(unable to translate for 2 seconds)…
7. I have, no I have sufficient…
8. but the position now might be at least equal.
9. Normally I would move the pawn on the kingside
10. trying to create an attack there
11. and the fact that his Q on b3 should help me
12. because the Q has gone a bit too far.
13. Now, also I don't have too many pieces to play with
14. so I can't try to do something on the kingside for that reason
15. I might consider moves like Qe5
16. then Ng4.
17. Em, another thing I might play is Ng4 straight way
18. because that eyes both f2 and e3.
19. That somewhere and move my pieces.
20. I might take the pawn on e3 in some variations
21. now I can consider immediately taking Bxe3
22. fxe3
23. Qxe3+.
24. Whether that is a good move…its hard to tell
25. but it does look sort of promising
26. Bxe3 is the most forcing of them
27. and white would have to take on e3 really
28. either immediately or after taking on c8.
29. he has this intermediate move.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

30. Ok lets start looking at the rook take first
31. so Bxe3
32. Rxc8
33. nothing there with the bishop yet
34. so I have to take back with the R Rxc8.
35. He has moves with the B such as Bxg6
36. but they are not really good
37. because I can always take back on f2
38. or p back on g6 I think…
39. So that's ok.
40. Bxe3
41. Rxc8
42. Rxc8
43. fxe3
44. Qxe3+,
45. Now I've got 2 pawns for the B
46. Rf2 is not good
47. because Ng4.
48. So he has to play Kh1
49. then I can but my Ng4,
50. and that leads to a very unpleasant threat of Nxf2,
51. and if in this position I can win his R for N I'll be a R and 2 pawns up.
52. And even if I drop one of the pawns I'm most likely to better in the position
53. because he…Bxe3.

S4 Position 3: W

1. So this is from the English opening
2. and the position looks about equal.
3. I would say it looks about equal
4. it's a real even
5. so not much seems to be happening.
6. I have a real problem with the Bb2
7. which seems to be blocked
8. so ideally I would like to play d4 at some point.
9. Simply to improve the B.
10. It may not be possible right now
11. but is definitely a plan.
12. Semi-open file
13. but black is defended that and reasonably well…
14. about playing, I can Nc4
15. whether that is a big improvement is not clear
16. but then I might take back with a piece
17. and then I might take back with a pawn b pawn…
18. slightly but that's not really an improvement
19. because the Nb6 is not doing much as such
20. so, black's position is pretty solid
21. so maybe I don't have any advantage (cough)
22. so, to have to have advantage is not to be sure in this position.
23. Black on the other hand doesn't really have a threat
24. Ne6 but that's not a threat as such,
25. em perhaps one way of playing would be Nf3
26. and then at some point if I can manage I might go for e4.
27. I have to watch out for the d4 pawn…
28. something has to be done with this B (Bh2).
29. Its just sitting there doing really nothing.
30. Now the other idea would be to Nc5 maybe via e4,
31. so I can probably start by playing Ne4

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

32. then moving the queen out
33. if back to c3
34. and then the Nc5…
35. not a big plan but its something.
36. So Ne4 (cough)
37. now e4 allows him to play Ne6
38. and after taking everything on b4
39. its nothing no I don't think it does.
40. The e5 pawn is hardly a weakness
41. and I don't think I can pressurise this
42. although Bc3…
43. (inaudible) a4
44. would be good enough for black.
45. Probably Ne4 I would play
46. with the idea of bringing Q out
47. then maybe Nc5.
48. Another option is to play Nf3
49. and watch out for eventual,
50. it could be equal.

S1 Position 3: W

1. Ah ok, we have a typical position from the English opening
2. white's fianchettoed.
3. A reversed Sicilian,
4. eh black has well a good,
5. actually white em should superficially seems to be quite well on the queenside
6. with pressure on the c file
7. with black a backward c pawn
8. fianchettoed B
9. em…and eh it seems that he's doing ok.
10. These positions are quite unclear
11. as is this because eh eh black has a very simple idea which is Ne6
12. followed by Qe3
13. followed by Qb5 / Q
14. Re8
15. Ne6
16. will help have a very strong influence across the board
17. discourage d4
18. em maybe in conjunction with a move like Qe7
19. give the possibility of Ng5
20. pressurising h3 pawn or h3 square
21. allowing the black Q in
22. some tactics eh…
23. its difficult to say.
24. White should probably try to break up the position,
25. liberate the B on B2.
26. there's two ways of doing this…
27. firstly try playing b4.
28. b4
29. is is eh is not such a hot idea
30. because of eh Na4 immediately
31. em nasty little trick probably NxB
32. Qx
33. Qb5
34. maybe pressurising b4
35. and eh Black is doing quite well.
36. He could play a4

37. but that would reduce really re…tension in this position,
38. and eh perhaps after Ne6
39. and an eventual a5
40. Nb5-b4
41. he may have trouble eh …bringing Q in …
42. ah that being said I wouldn't be hasty about playing that move of
43. …moves pawns
44. with the knight into e4
45. which could pressurise the d6 square
46. which could be quite vulnerable…
47. I don't like a4 though
48. because its very committal
49. and eh its losing any chance of taking in the future.
50. I'd problem with e3 is that after Ne6
51. eh a move like Qc2
52. Qd5
53. followed by Rd8
54. Rd7
55. then e3 becomes weak.
56. So, what has actually black
57. or eh white's best course of action
58. is simply tricky to assess…
59. and indeed I'd go back to the move I didn't like earlier which is a4
60. is not so stupid as it looks
61. followed by Ne6
62. Qb5
63. after an exchange of queens
64. white has gained…
65. eh I mean white will have doubled pawns
66. but the a file will be opened
67. Nc4 attack
68. a5
69. and if Nxc4
70. take with the b pawn
71. or if white so wanted
72. although I more likely to take with (the) a pawn…
73. the idea of stinging the pawn.

S8 Position 3:B

1. Again a complex position.
2. Open basically.
3. Now lets have a look.
4. Material level again.
5. The king is fianchettoed.
6. Ah hmmm, lets have a look,
7. what do we see in this position?…
8. I think black has rather a nice position.
9. R on seventh.
10. Bb2 doesn't do anything.
11. Em ok b4-b5 might be an idea for white.
12. Or black can double on the d file
13. but then how do you proceed?
14. Ok move Q
15. with Rd5
16. and then at some stage, yea maybe, ah that's an idea yea, ok double on d file
17. play Nc8
18. Nd8

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

19. then e4 later on.
20. Do you have to take care of anything?
21. No I don't think so,
22. em alternatively well developed black,
23. f5
24. e4 is also possible
25. but that would make Bb2 alive again.
26. Rad8
27. and Rd7
28. so Rd7
29. I might prefer Ne4-c5
30. could be a problem
31. so maybe Rd8 after all
32. with idea of playing Nd4
33. attacking e2
34. might weaken e3
35. might be weak
36. oh that is better.
37. c7 is covered,
38. a5 little problem,
39. easily just sac the pawn.
40. So, Rad8
41. looks like the logical move now…
42. Rad8 is my move.

S17 Position 3: B

1. Ok, em…hah, I know what opening and everything this came from.
2. Its an English…
3. em…yea I hate these positions as black
4. because em…they're just equal probably
5. em except they're far easier for white to play…
6. this position…solid pawn structures
7. white would have the edge
8. because white has an extra central pawn.
9. White has a Bishop…
10. em black wants to use the open d file…
11. and play things like Nd4
12. to stunt the bishop
13. but its very…lots of organising of
14. the c pawn would be often un pres in those lines.
15. Rook can't move
16. as protecting (the) a pawn.
17. Candidate moves…
18. I would be looking at what white might play
19. b4 in that position…
20. no white can't play b4
21. as I reply Na4.
22. Em…then white might play moves like Nc4
23. to get rid of that half decent Nb6.
24. Em…because the Nd2 would be a much worse piece
25. so he might just exchange it.
26. So Nc4
27. move my Nd7
28. Qb5
29. but my Nb6 and pawn b7 vulnerable
30. so perhaps just exchange…
31. again play Qd6 here

32. because the Q looks like its half decent on that square.
33. Or play Rd7 here
34. with the idea of playing Nd5
35. then d4.
36. Actually that looks like a more sensible suggestion.
37. Rd7…
38. coaxes the idea of playing Rd5
39. em…it attacks the a pawn
40. and after that bring the rook over.
41. Eh, allowing me just to play Nd4.
42. There always a Q +,
43. and actually Rd7 allows white to play it now.
44. Na4
45. will no longer work because of Qc4+.
46. So Rd2-d4
47. is not the biggest problem in the world
48. but …just reply to it by…well reply a few ways maybe
49. a4
50. would be the best way eh,
51. but he can also play Rd4
52. try to stop…the Nd4.
53. Ah…oh yea the other idea is that the white Kg2…

S3 Position 3:W

1. Definitely looks better for white.
2. My pieces are considerably more active than his,
3. my,
4. for the moment he's holding his weakness on c7,
5. em the real question is how I can bring a little more pressure down on his position.
6. Em pa5 is a little bit protruding.
7. Bc3 looks reasonable tempting.
8. He can play a4 he might,
9. d4,
10. a4 immediately
11. could be the the move,
12. em because Bc3,
13. I don't seem to be able to do because Nd7 is possible
14. Qxc7,
15. so he would try a4…
16. then c6
17. Bc3
18. Nd7 in that position
19. which would force my Q from c5,
20. so other than that its difficult to get anything going in the centre
21. or kingside
22. because f4
23. would leave my e2 pawn undefended
24. and my pieces aren't posted there anyway.
25. I would like to make more use of my B,
26. em it doesn't seem that prosperous at the moment.
27. Ne6 is coming up next move as well
28. so it will be quite difficult to keep Qc5,
29. eh …so if N,
30. a4
31. Ne6
32. Qb5 is interesting
33. 'cause if he takes px

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

34. he can't get at my pawn weaknesses I don't think.
35. I've got a lot of activity in that position.
36. a4 is definitely looking like the move, em.
37. Other than that I'm not sure.
38. I'd like to make some use my Ra1.
39. I really don't see how that possible for the moment.
40. Em, so in that case a4
41. I think definitely strongest move here.
42. Ne4 but that's
43. Ne4
44. Ne6
45. Qe3
46. Nd5
47. I don't think I'd like that.
48. Yea definitely a4…
49. Ne6
50. then Qb5
51. can't play c6,
52. N undefended.
53. a4 would be good I think…

S13 Position 1: W

1. Em, I'm thinking about playing Bxd5,
2. then cxd5 em,
3. maybe Rf3
4. to way off black's r
5. and threatening to check on the c file.
6. Em, eh I think an immediate f5,
7. seems to lose the exchange after exf5
8. gxf5
9. Bh5…em,
10. do do I'm thinking about Bxd5
11. cxd5,
12. em Rf3
13. Rxf3
14. Qxf3,
15. maybe Bd4,
16. em Qc3+
17. Kb8 something like that,
18. em I'm thinking about playing Qg2…
19. and to move to chase the rook.
20. I don't see any threats for the N
21. or the B.
22. I don't see any way he can defend the Rh3.
23. Em, I don't see any good square along that rank for it to move to,
24. so perhaps defend with other R.
25. Qg2
26. Rd8-h8
27. Bxd5
28. cxd5…
29. maybe Rh1
30. looks quite even.
31. I'm thinking about a4
32. to be playing a5
33. and attacking Qb6
34. or maybe Bxd5
35. cxd5

36. and then a4
37. or maybe and then Bb4.
38. So, Bxd5,
39. cxd5
40. Bb4
41. view to going to d6
42. where the Bishop is very well placed,
43. em I don't see a good defense for black to that at all,
44. ok so Bxd5
45. cxd5
46. Bb4
47. maybe he plays Be4…

S20 Position 1: W

1. I'm white again and…
2. I do think generally speaking white would appear to have the best of the play on the board…
3. he has more scope for his pieces.
4. The one hole in this and I think that is Bg6 on that diagonal
5. is a threat in front of the king
6. and eh,… I imagine we'll have to try to do something with this.
7. I don't see that there is any immediate threat.
8. I could of course play pawn to f5
9. which would lock off that Bg6
10. and px
11. px back
12. threatening the Bishop
13. and eh forcing it back,
14. and if those two advanced pawns might be good
15. but they're up against three
16. so I would be a little concerned about doing away with that,
17. I think the most attacking move would be with the Qg7 sorry Qg2
18. attacking the R.
19. Now has he any counter play with that?
20. The Nd5 could be threatening
21. Qe1
22. Qb5 has to be watched,
23. so but I don't think there's any threat of it.
24. The N even if it moves exposing the B to attack by Rd8.
25. That Bishop is still defended by the R on d…
26. So I think probably the Qg2 might be best.
27. I have to guard here too against a possible sacrifice of the Rh3,
28. Rxa3
29. and opening up a gap on the queenside there where the N,
30. or where his Qb6 has an attack,
31. so I wouldn't.
32. But I think it can be sustained,
33. so I think, all things considered I would play Qg2.

S12 Position 1:B

1. Ok, I'll have a quick look at the position here.
2. Eh 3-6 pawns for me
3. 6 pawns for white,
4. N and B versus B and B,
5. so material is level
6. em Rh8,

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

7. I've a nice open file here.
8. B attacking c2 eh,
9. none of my major pieces are under attack,
10. 3 pawns,
11. white has 3 pawns in the centre,
12. or on the queenside rather that are coming through
13. looking good,
14. eh N is being attacked on the open file,
15. potential for an open file, on d file,
16. so Rd8
17. looks like a good move,
18. em white's white B lined up against my king's corner
19. BxN
20. pxB,
21. Bh5
22. or Ba5 I should say
23. is a potential move.
24. If I don't actually move Rd8 em,
25. need to be aware of that,
26. no danger there at the moment
27. 'cause queen will just take,
28. eh strong pawns coming through on q side.
29. Attacking moves going forward,
30. N move.
31. Nxp
32. well protected by two pieces.
33. Eh N…
34. g4 pawn is over protected by the queen.
35. Eh strong Nd5,
36. I think my move would be Rd8. Yea…

S11 Position 1: B

1. Well, hits me as rather a complicated position.
2. I know that I am equal on material,
3. black (white) ahs two bishops,
4. immediate way to make progress.
5. Actually I have an open h file,
6. black (white) has a chance possibly to eh to advance f pawn …,
7. increase pressure that way.
8. My Ne5 (d5) well supported
9. but don't see any options for it,
10. em I think I might actually, possibly bring Ne7
11. and eh perhaps get some pressure,
12. if I bring it right back Ne7
13. also helps to helps to prevent f5
14. hold that up.
15. Em, I think eh that might be the best move to make.
16. I'm looking at my B isn't doing very much where it is (Bg6).
17. I don't see how I can improve it
18. eh I could possibly I think try and double rooks on h file
19. maybeRh7
20. then Rd8-h8,
21. white could try and swap off those rooks
22. but I don't think I would mind that too much,
23. I would still control the h file,
24. of course, em…em I think that is probably the best way to go about it.
25. I don't think my king is in any danger either.

26. And I don't see any way of getting at the em white king.
27. I'm particularly sure my choice would be between…

S2 Position 1: W

1. Ok, …first impression I get from this position is that its pretty even.
2. I'm looking at the eh, rook being uncovered by the N.
3. If black moves his N
4. the R is bearing down on the position.
5. But, for my move I think…I think I would probably try to get my b down in front of my pawn blocking the eh…R,
6. unfortunately the knight is covering the square which would be nice to get my B on…
7. b4 em,
8. however, if he takes that I have a check
9. with my RxR
10. and then moving my own R+
11. if he takes with either the king or queen.
12. Eh, so I think that would be the, the lines I would be conceiving along,
13. Bb4
14. if he moves the pawn up against the Bg6 (f5)
15. e x pawn.
16. The R on the site doesn't worry me too much unduly,
17. black's rook,
18. so I think my move would be Bb4
19. opening up my rook threatening something to happen,
20. he can decide whether to swap the N for B,
21. and eh give me an open file there.
22. I can retake
23. but lose a pawn,
24. but I think in a game I would do that
25. because time is just as valuable as the piece.
26. Is that ok?

S2 Position 2: W

1. Right, now first examination
2. again it looks like I don't have any particular advantages.
3. Em the material is pretty equal.
4. A queen two rooks and 2 pieces,
5. bishop and knight,
6. black has eh a pawn in the centre.
7. It looks good
8. but its also isolated
9. and there's little threats on here that we may not fully comprehend
10. like eh at some stage the bishop taking the rook
11. taking on Bc5 bringing one of the rooks to the back rank.
12. So, the only piece that's really unprotected so far, the one pawn on the right hand side of the board on h6 for black
13. but I don't think its very weak as such
14. em…but if I move my Nf4
15. threatening his queen
16. threatening the centre pawn
17. and threatening the pawn as well on g6,
18. you know, its going to be dangerous,
19. especially if I worked out the sacrifice on taking the g6 pawn
20. might be good,
21. because my queen is on the same diagonal as his king

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

22. and em it looks like pressure.
23. So the knight at the moment is doing nothing
24. except protecting d4 should he advance the pawn
25. offering to swap queens
26. and if I move Nd4
27. he could Bxd4
28. and after RxRc8
29. Rxc8
30. his rook would be on c8
31. controlling that file
32. so my move at the moment I'm thinking of is immediately Nf4
33. which apart form other considerations threatens his queen
34. and he has to move it.
35. So, and where he moves it to I would wait to see
36. but eh obviously can't move it away from the pawn
37. or it falls by the queen
38. so I'd be happy enough with that move if I were to play this position as white.
39. That 3 minutes? Yep.

S20 Position 2: W

1. Ok, right. Now first of all suppose its as well to see how we're balanced on material…
2. we're pretty well equal,
3. so starting from scratch here,
4. and eh, he has , black has some threats here.
5. Again it's a question of whether its worth sacrificing
6. if Bc4 (Bc5) was to take the pe3,
7. pawn takes back
8. fxe3
9. then Qxe3+
10. and is checking
11. and at the same time the file where the Bc was I the c file is opened up to black's Rc8.
12. Now the alternative is that white Bxe3
13. RxR,
14. but the other rook,
15. that Rf8 would take Rf8xc8
16. and hold that open file
17. and then the Qxe3
18. could be very serious.
19. I think from white's point of view that's the most immediate threat,
20. so how to answer that is the question for white.
21. How for white to try and advance himself?
22. What has he got? And they say attack is often the best form of defense.
23. Would be Ne2-f4,
24. that threatens the Qe6,
25. it doesn't' really threaten the pawn black pawn on d5
26. because it has the queen and knight covering it.
27. But, the queen has to move.
28. And the question is where does the queen have to move to.
29. Em, … the queen might consider moving to e5.
30. now there's an inherent threat in this on Qe5
31. if black plays Ng5.
32. there's a threat of mate on h2
33. but it can be defended by g3,
34. so I don't think I'd worry too much about it.
35. I'd see, where would I move that queen,
36. or what would be my follow on to there with that threat.
37. Eh…took…

38. I think white eh eh…is eh…his pawns are pretty strong.
39. Black has the isolated pawn d5.

S11 Position 2: W

1. Well, the eh, first thing I notice about the position is that ah material once again appears to be level.
2. However, in this situation em first thing that comes to mind is the fact that black has a weak, potentially weak pd5.
3. it can't really advance I think without being lost
4. and possibly eh I might have some opportunity of ganging up on it,
5. em maybe by playing Rf1-d1
6. and pressurising it that way,
7. em also I notice that my own king seems to be pretty safe.
8. There are no threats against it at the moment.
9. And eh there are, some of blacks pieces are not in very favourable positions.
10. The queen for example can be attacked by Nf4,
11. which also helps put pressure on the pd5.
12. so maybe some move in conjunction with Rfd1
13. followed by Nf4.
14. don't know which one of them I'd pick first.
15. And also em, there's another weakness in black's position.
16. I don't see any weaknesses in my own position.
17. Em he could try and swap off the pd5 by advancing it maybe
18. if I exchanged queens
19. but then I can just recapture
20. and he's swapping one weakness for another
21. because immediately if he swaps off queens (after d4)
22. he captures with pawn
23. then the pg6 is weak as well.
24. And the other factor of the position is ph6 is weak
25. and eh, so I might have some possibilities there,
26. but Nf4 as well as attacking the queen
27. would also, em look at pg6 as well
28. and maybe some pressure around it there
29. Qc2 and possibly attack,
30. attack there so that seems to be a weakness.
31. Em the only underdeveloped piece that I see that I have is Rf1
32. so Rd1…

S12 Position 2: B

1. Again lots of pieces on the board
2. eh pawns 6 versus 6.
3. Two bishops,
4. two knights on the board,
5. two rooks,
6. two queens
7. so material is level
8. white has advanced b5 pawn
9. supported by queen and bishop.
10. White's rook is attacking my knight or my bishop.
11. Bc5 em by the Nf6 protected by the queen.
12. White doesn't have a black square bishop to attack it
13. so, that's ok, by queen
14. or knight.
15. Eh, immediate danger to me
16. queen is supporting a pawn

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

17. which is protected twice.
18. N move f4 (Nf4)
19. hit my queen make that move.
20. Rd1 from f file for white
21. looks like potential good move.
22. Em, that would put more pressure on my d pawn
23. which is isolated.
24. Probably reinforce the defense of that by potential Rfd8.
25. White pb5 could that move forward to b6
26. but do I want it to go to b6?
27. so by playing b6
28. that would stop that advancing a bit more.
29. Em put my pawns on black squares
30. potentially good as white has a white square bishop
31. so b6 looks good
32. that also adds more defence to Bc5,
33. eh …so two crucial things here
34. pd5
35. and Bc5
36. protect them I think
37. I would play Rfd8. Yea…

S13 Position 2:W

1. Ok, I'm noticing Bc5,
2. em I'm checking for material equality,
3. it seems that material is equal.
4. I see that black has an isolated d pawn,
5. em I'm considering sacrificing on g6
6. although that's very premature at this stage.
7. I'm looking at Nf4
8. which attacks the queen
9. and the isolated queen's pawn at the same time.
10. Nf4
11. Qd7 possible,
12. of so Qe7
13. bad because Bxg6
14. pxg6
15. Nxg6
16. with fork on queen and rook
17. so ok, Nf4
18. Qd7
19. and perhaps Rfd1
20. in that case to put more indirect pressure on the pawn (IQP).
21. Or perhaps the manoeuvre Be2-f3
22. or perhaps the idea to blockade the d4 square in that case.
23. So I'm looking at Nd4
24. Qd7
25. and Qb2
26. then attacking knight,
27. perhaps if black can just play Kg7
28. defending knight
29. and then maybe I can play Rcd1
30. to double,
31. indirect attack from the rook and the pawn.
32. Looking at, I don't see any threats for black
33. and my king is em.
34. I think f4,

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

35. that's bad because it loses pe3.
36. Bb1
37. and then Rfd1
38. with good control over the d4 square
39. but that might allow black to immediately play d4.
40. Bb1
41. d4
42. can't take 'cause Ne2 is undefended.
43. So that allows him to liquidate his isolated pawn.
44. As perhaps I play Rfd1 immediately…

S13 Position 3: W

1. Em, I've noticed that my bishop (Bb2) is blocked by the pawn chain e5,
2. I'm thinking about playing Ne4
3. and with the possibility of maybe sacrificing on f6.
4. I'm thinking about playing Qe3
5. em with the idea of transferring my queen to kingside.
6. Although all my pieces are on the queen side
7. so perhaps I should be playing over there.
8. That's what well against my instincts.
9. I'm thinking about Nc4
10. Nxc4
11. dxc4
12. and then double rooks on d file
13. or perhaps Rc2
14. followed by Rac1
15. and try and gang up on thee c7 pawn
16. which looks weak.
17. I'd say that's the main weakness in black's position.
18. And eh I see his Nd8 is passively placed.
19. I'm thinking about Nc4
20. Nxc4
21. Qxc4
22. pinning rook,
23. I then playing Qh4
24. and maybe transferring some pieces to the kingside.
25. I'm thinking about playing em Rh1
26. with the idea of h4-h5 with the kingside
27. and maybe attacking there,
28. em should move e4
29. but that creates a long-term weakness on e3,
30. do something about offering controlling d5.
31. I see that black is ready to play Ne6
32. winning a tempo on my queen,
33. so I don't think Nc4
34. would be such a good idea.
35. Em, so I'm inclined to move my queen Qe3 I think,
36. possibly Qe3
37. Ne6
38. a4
39. with the the idea of playing Ba3 where it stands well,
40. small weakness of d3
41. em, I'm trying to see f5
42. looks to be a good square for my knight
43. so perhaps I would play Nf3
44. Nh4
45. Nf5

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

46. a4
47. in conjunction with a3 Ba3,
48. attacking dark squares…

S12 Position 3: W

1. Ok, lots more pieces this time,
2. 3-7 pawns,
3. knight and bishop,
4. black has 7 pawns
5. two knights versus bishop and knight.
6. My queen is in very open play.
7. Looking a bit tied up at the back
8. although my Rc1 is in a good position.
9. My king is on a white square
10. and no white square bishop for black.
11. That's ok.
12. I'd be worried if he had a light squared bishop.
13. Eh I'm attacking the pc7.
14. Eh, with my queen,
15. and I'm attacking it twice with queen and rook.
16. Looking over towards the centre eh squares my knight can go to my knight has control of e4/c4
17. look like nice squares for it to go to.
18. On the defensive side black isn't particularly aggressive at the moment.
19. Nb6 can come in to d4 d5 I should say
20. to not great advantage
21. be lost if it went to there,
22. em a5 pawn not doing very much,
23. attacking it with the queen.
24. I could hit it with the bishop by going Bc3.
25. Eh, play then pawn up
26. no real in that.
27. My knight is protecting c4,
28. can the queen be attacked if I play,
29. black could play Ne6 and hit my queen
30. and protect the c7 pawn.
31. Queen back to e3 eh
32. stay on the c file? …
33. just looking again queen move back along c file
34. or diagonally across?
35. Might play Ne4
36. and that is looking like the move I would make
37. Ne4.
38. Eh I can get , gets hit by f5,
39. Qxe5
40. I can do that
41. or Bxe5
42. pxe5
43. so I could put pressure on the e5 pawn…by playing Nf3
44. and I don't think that does anything positive for…
45. I think in this position I would play Nc4…

S11 Position 3: B

1. Well, this is a normal position, thanks be to God!
2. …em I eh, see that my king is quite safe.
3. Em, the knights look possibly em threatening at some stage to go to good squares.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

4. I eh black's white's eh some of white's pieces don't seem to be very well placed
5. eh his rook two rooks, his queen and rook lined up on the c file
6. eh don't seem to be very threatening
7. and his Ra1 is not doing very much.
8. Em I can't see any immediate threats,
9. but the knights and possibly…
10. I think maybe move may own rook
11. and probably play, either of my rooks
12. even both over to d file,
13. em yes and em I don't see any eh, I don't see much scope for my knights at the moment
14. eh, eh my queen is on e8,
15. possibly queen,
16. no rook Rf7-d7
17. with the prospect of maybe doubling them
18. and getting queen out somewhere
19. and get a threat on the king.
20. Position wise I don't see any serious weaknesses for either team, either side.
21. Eh looking at pawns so if
22. em, as I say his bishop is basically, on the long diagonal,
23. is em …not actually threatening very much the em
24. and also the queen and rook lined up,
25. I think eh I could possibly play Rd7
26. with the possibility of doubling up. Rd7-d5,
27. or even the Nb6-d4,
28. and maybe some pressure going em…

S20 Position 3: B

1. Well, I'm ok I'm ready to go.
2. I'm looking at black first of all this time
3. and eh, so far as the minor pieces are concerning the rest is equal
4. but its two knights against a knight and bishop.
5. Now oddly enough I'm the sort of guy who likes to play with knights
6. so, I think I have possibilities here.
7. There's nothing to be won immediately
8. but as well as that his Qc5 is pretty well exposed out there.
9. And I think the three pawns on the queenside should bring somewhere
10. but it's a very very equal position.
11. I think as a first off goal what I would be looking at would be probably Nb6-d7 (I corrected the notation here, he had said Nb3-d2 by mistake)
12. attacking the queen
13. that forces a move of some sort from the queen.
14. And I think that would be a starting point
15. because then you could get the pawns that are behind the knights to move.
16. Now queen can't take c7
17. because its covered by the pawn
18. and would have to move pretty well back away from the influence of the two knights
19. which cover almost up as far as the fifth rank,
20. the way they're placed.
21. So, the queen back so (Qe3 or Qc3?)
22. and probably, neither of them appear to be great.
23. He can go to eh…he can go to e6 (Qe3)
24. probably the best move.
25. On the other hand, if he could move back to c6 or c7 (c3 or c2)
26. the queen is pretty locked in there.
27. Back to Qc3.
28. there's a move with the knight,
29. not yet!

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

30. But, eventually…eh there would be Nd4
31. threatening the queen again there.
32. Eh black, as I see it would have to start the moving those knights to get value out of them.
33. So my choice move would be Nb3-d2 (Nb6-d7).

S2 Position 3: W

1. The thing open is to be to try and force a move in the centre,
2. so I would probably be playing something like Nb4
3. or d4
4. because the bishop obviously has to get more freedom along that line…
5. eh the other thing would be going on to f4
6. but unfortunately I think that would be giving black too much.
7. If he took the pawn on f4 with his e5 pawn
8. his queen is then ranging down on e2 for an easy check…
9. em eh the rook is in front of the king (on f7)
10. and unfortunately I don't have a white bishop to pin that somewhere
11. and the one obvious move of getting down on c7 (Qxc7)
12. but its covered by that rook.
13. Even though I could take but I'd lose material.
14. So…I'm thinking at the moment that my best move is to play d4
15. even allowing black to retake
16. or to move ahead with it…
17. eh I just think it gives me a little bit more freedom
18. and if I could at one stage shift the the rook off by a knight move or something
19. I would looking dangerously down to c7
20. with queen protected by my rook
21. and there's always the possibility of some sort of dramatic take by the bishop
22. eh again as its lined up against the king
23. although I can't see it working at the moment
24. but its always the the type of thing that worries,
25. would worry me
26. eh like the bishop continually focussing down on g7 in front of the king.
27. So as to what move I would actually make it would be in the centre
28. d3-d4 pawn
29. or the knight form d2-e4
30. or even Nc4
31. and even the back pawn e4
32. forcing the pawn to stay where it is until I can attack it on the other side…

Segmented protocol transcripts for random positions

S3: Position 4: W
1. It's a mess, em …
2. yea eh well just look Rc5 briefly.
3. The position is quite difficult to make sense of.
4. Em ok basically both kings are in massive danger.
5. Well maybe black's king move
6. because black doesn't have a check.
7. If I could deflect his Qd5
8. and Rf2
9. from defended f3 pawn,
10. then Nxf3++.
11. Em so the question is… how do I manage to do that?
12. Em I've a feel like I should be doing something else,
13. but maybe mate I is the only thing to play for in this position.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

14. Material it doesn't matter.
15. But, I don't really see any other way.
16. Right eh ok, so lets see Rc5
17. isn't possible because Bxc5 em,
18. although I could take on c5 with queen
19. then queen takes queen possible
20. deflect queen from f3,
21. deflecting rook as well more difficult.
22. Eh so, yea… Rc5
23. how can I do that.
24. My king can't move.
25. My knight for the moment can't move.
26. Qe1 idea of threatening R.
27. Qe1 is probably one of the more promising ones.
28. Em just trying to get rid of his Rf2
29. em give me option of playing Qd3
30. which in some ways might be useful.
31. Like after Qe1
32. Rg2
33. Re8
34. Bf5
35. some combination of Qh5.
36. Yea if I can get his queen off that square.
37. Em, ok eh Qe1… eh Qe1 I think does,
38. I might just look Rg2,
39. I'd imagine that's quite weak.
40. Em lets see if I have any ideas.
41. Ok once again if I can deflect the queen em.
42. How do I deflect the queen though?
43. Qe4
44. Qx
45. Rx
46. then Rf2
47. can't ,
48. Ne5.
49. I definitively have to proceed with forcing moves.
50. Once he gets Ne5
51. he's better
52. although I don't know how anyone could be better in this position.

S17 Position 4:B

1. Ok we have a complicated position.
2. Em, where it looks as though …it wouldn't surprise me as black, eh, if I was to be mated in the next few moves.
3. So, em I just better count the material.
4. 1-2-3-4-5-6, ok 1-2-3-4-5-6 (counting pawns).
5. Ok we're even on material.
6. Em, white has a whole load of pieces around my king
7. and em I don't like that.
8. Em, (laughing) and really in this position I would be looking at eh like I suppose I do have some decent moves.
9. Em eh things in my position.
10. No I don't! eh (laughing).
11. Well actually I don't really have anything significant.
12. I'm playing for a draw here.
13. Eh eh, I have to try and find a way eh of white's threats are
14. and countering them.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

15. Now …white would like to play…oh white would like to get his bishop
16. and queen involved
17. Qe5 would be a decent move for white,
18. em because its just threatening Rxa7+
19. eh Rxa7+
20. Kxa7
21. Qc7+ mating him
22. and actually that looks extremely difficult to avoid for white.
23. So, do I play a move like Rf8
24. to try just to stop that
25. em perhaps I have to
26. em do I have any other nice ideas if I play,
27. I suppose I might play Qe8
28. after Qb5 I might be able to get away with
29. Nc6,
30. so harassing.
31. Eh oh hold on does Rf4,
32. no there isn't Rf4,
33. em because the Bc1.
34. Is there any way I can make that tactic work?
35. No there isn't,
36. eh my king doesn't have a move,
37. my knight doesn't have a move.
38. And so lets just look at the pieces I can actually move in this position.
39. Em, Rf4
40. just the idea of Nd4,
41. looks really strong for em.
42. Oh perhaps it doesn't
43. Nd4
44. RxR,
45. Nx…
46. interesting.
47. Lets just see.
48. Eh, can I play Qd8.
49. Qd8!
50. Qd8 attacks the rook em.
51. Qe5 no longer has the same threats
52. because Queen covers the square that's going to
53. ok Qd8 looks like a decent thing,
54. and on the following move I would play something like Rd1
55. so that my bishop on f1 in protecting that pawn on b5
56. and I can try and get that bishop back into the game.
57. Oh hang on maybe I should just play Rd1 immediately?
58. Because that threatens Rxc1
59. Nxc1
60. Rf4
61. and if white plays Qe5
62. I have my defence against that of
63. Qd8.
64. not that I'm entirely happy with that.

S4 Position 4: B

1. Well, here some kids have been messing again,
2. and there have another random position on the board.
3. At least this one is good for me.
4. My king looks safer than the enemy.
5. And or does it?

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

6. Yes it does.
7. Material seems to be…equal.
8. Right material is equal
9. and I hope to bring my bishop.
10. The white king is not a very powerful creature.
11. My king although more in the centre is the safer one
12. Rc5 is not a threat
13. because I control that square.
14. Nxf3 is not a threat at all.
15. I would be happy to see that (Qxf3++).
16. Now can get at his king
17. that would be pretty nice thing to do.
18. So how about h5
19. so I
20. and well checkmate and everything.
21. So he cannot take
22. because Qf5 ++ as well.
23. So that, he would have to move the rook somewhere to stop that.
24. Which sure enough gives me an idea really
25. maybe I should start with f5
26. because that doesn't give him the same chance
27. but f5 exposes my king a little bit,
28. he ok
29. so if I whatever reason would be better,
30. also if I can now get my other rook on a7 to h-file
31. now that would be checkmate immediately.
32. h5
33. he moves the rook somewhere
34. I might move Rh7
35. if the bishop goes
36. I take on g4.
37. Bxg4
38. my bishop comes somewhere say Be6
39. and rook-h
40. and checkmate is coming.
41. Oh I would say that is the move,
42. what else what else have I no no yea h5
43. looks the best
44. and I don't see other,
45. ok if I can get my Qh4
46. that is checkmate too.
47. Some one idea is to play Kh6
48. where the cute idea of Qg5
49. Qh4++.
50. Now how he is going to defend that I don't know.
51. Maybe, how about this, how would have to take off Rxf6
52. with rook
53. and then if I play Qg5…

S8 Position 4: B

1. Hmm, this is Alice in Wonderland chess (laughs) oh gosh, how did this position arise? Eh ok this could only be a position out of a psychological test!
2. Em yea takes time to understand this.
3. Wonder if this position is possible?
4. Ok so black to move.
5. Both kings are in danger hmm.
6. Ok so, what are the features of the position.
7. Right probably can only be mate,

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

8. em better look at the material 2-4-6, 1-4-6 (counting pawns),
9. equal material
10. but that's not very important in this position.
11. God and I'm expected to come up with a move.
12. Oh, ok so how do you get at the king?
13. Hmm, Rf4
14. doesn't work
15. Bc1 can take that.
16. Nb3 covers a5.
17. It's a tricky position.
18. So ok the Rc7 not protected.
19. What can we do there?
20. Em I'm supposed to talk all the time.
21. I prefer to think in this position (laughs).
22. Ok can't find a sensible move.
23. Glad I'm not in time trouble.
24. Ok so its, if I look just logically here…
25. Rd1 any sensible move,
26. trying to get rid of that Bc1
27. then Rf4.
28. But after Rd1
29. Be3
30. hasn't really achieved anything.
31. On the other hand, ok is black really in danger of being in mate here?
32. Could be, hmm.
33. Ok if I come up with a move its because I have to not because I really like this position.
34. Em these are not my type of positions.
35. Ok Nc6
36. that would also be something
37. so Qd8
38. no
39. then Rd7
40. sacking the exchange.
41. Qa5 is covered
42. hmm very difficult em you will probably tell me that my time is up soon.
43. Ok, Rf5 is also a possibility,
44. don't ask me why,
45. hmm probably Rd1 looks most sensible
46. Be3 than Re1
47. but it doesn't really do anything.
48. At least it covers b5 with the bf4.

S1 Position 4: W

1. Well, first this position strikes me as being eh, composition straight off
2. or eh a random position
3. due to the placement of all the pawns everywhere.
4. The f pawns the 'Irish pawn centre'.
5. Also the eh…the way eh the pieces the pawns are all just the white pawns are all just atypically spread.
6. The actual position to assess, phew em quite tricky
7. because the placement of the pieces suggest that both teams are in danger,
8. they actually look quite vulnerable
9. it seems to me that the white king is in more danger…
10. eh than the black so to speak.
11. Em obviously eh obviously eh white would, white would like to play Rc5 at some point
12. eh taking advantage of the queen and king unfortunate alignment of the queen.
13. The black pieces seem more well placed.
14. The material here in the position is quite irrelevant actually

15. from eh looking at it its actually…equal.
16. Eh the knight of course would be more useful in this kind of position
17. bishop should,
18. tends to have a more limited role.
19. So the bishop when you scale it sown essentially worth nothing
20. protected the Re6,
21. now , plans in this position might be eh…moves like f5.
22. it is white to play.
23. I need to support, to assess the kind of threats black has
24. as I said I don't see any immediate eh chances of there being eh an advantage for white.
25. That being said the Ng1 could spring into action at some point,
26. the Rf2 seems to gain control,
27. covering it as regards seeing what eh white should do in this position its quite tricky to say the least …(keep on talking)
28. essentially its going to be difficult for white by
29. I don't see eh the winning path for black either
30. but eh all the activity suggests there maybe moves like h5
31. gxh5
32. then gxh6
33. which immediately might just win the house straight out. That eh…

S2 Position 4: W

1. Well this looks like what I would call a hairy position.
2. The king is out on the edge of the board (Kh3).
3. Risky position indeed.
4. And the black pawns are at very advanced position.
5. So, I , I'm looking in consequence at the black king
6. and I see that he is also in a very dangerous position
7. because he can't actually move the king anywhere (Kg5)
8. and he will be checked because on every square he could move to he would be walking into check.
9. So, if I could force the issue like swinging the Rc5
10. planting it in front of the queen.
11. Unfortunately its covered by the bishop (Bf8).
12. So it would be Bxc5.
13. the other alternative would be to play Nxf3+,
14. getting the queen out of the way first,
15. but unfortunately that would give black the game
16. because Qxf3 mate.
17. So, very delicately poised…
18. em another possibility would be to try bringing the rook over.
19. That doesn't seem to work.
20. So what would I move, I … the key lies somewhere in shifting the queen off the file…
21. off the rank in front of the king
22. so I get the check in with the rook.
23. So, can I take the bishop first? (E: keep on talking)
24. ok, its all hinging around there so maybe if I could … no move.
25. So, supposing I play my rook in front of queen…

S20 Position 4: B

1. This is certainly a most unusual position for both white and black with both kings. The position their in their absolutely the game is wild.
2. I couldn't describe it as anything else.
3. Eh: Black has queen and rook on a file (rank)
4. and a very far advanced pawn
5. which I don't think will be tenable for some time.

6. The black king and the white king are in most disastrous looking positions.
7. So, what would I do as black.
8. Well I think obviously.
9. One is going to be attempting to force the position onto the white king.
10. Now the difficulty, again even the Rd3 is defended by the Bf1.
11. So white or black with the move at the moment could I think immediately play Qd8. now that looks like the most immediate response.
12. It threatens the rook from c7.
13. the white rook which has really eh, no where to go.
14. The white pawns of course can be advancing .
15. But I think this is the move, now play that and eh.
16. The Rb7 in which,
17. ah there's a knight up there too
18. which makes life rather difficult
19. because there was a lovely move with the queen,
20. but however, Qa5
21. would have been very nice
22. except there's a knight there so you can't do that.
23. Its very hard to see how one makes progress in this position.
24. But I think maybe on second thoughts the Qe8 might be better than Qd8
25. now this threatens the pawn
26. if there is a way of shifting something there
27. but where its going to shift to I just do not know.
28. Is there anything else in this at all… a most confusing position.

S11 Position 4: W

1. Well the first thing that occurs to me when I look at this is that it is a mirror image of the position,
2. eh, take me a moment or two to realise what colour I am.
3. I think I'm white (laughs) (E: yes that's right).
4. Em ok, I em, well first thing is my king is in a very exposed position
5. but em maybe eh black has some threats against that,
6. em my Re3 is defended by bishop at moment.
7. Maybe black could generate threats against that
8. eh by playing Rh1
9. and maybe getting the rook over,
10. well Rh1
11. threatens RxB
12. the QxR.
13. Also maybe trying something, his rook over to the em h-file
14. his king is his king is in his way but I don't see the ,
15. any immediate yes, does it? There's a threat there.
16. I'd better do something about it. I could, I'd like to be able to play Rf4 (c5)
17. that would win the queen
18. but I can't at the moment because of Bf8…
19. em, I don't see any maybe if I played em, if I play Re1 (Re8)
20. the double threat of swiping off the Bf8
21. and then playing Rf4,
22. that's a major threat,
23. so what happens if I play Re1 (Re8),
24. eh well, if bishop moves to g7…

S12 Position 4: B

1. Ok, I'm black in this position,
2. piece count 2-5-6 pawns.
3. White has six pawns two bishops, two knights (for black),

4. so material level.
5. White has rook on seventh (Rc7).
6. Two passed pawns.
7. White's kings exposed on b4,
8. can't really move so have the potential in there for checkmate playing something like Nxc6+
9. but that loses to Qxc6++ eh,
10. I can pin the queen by playing Rf4
11. that wins the queen as
12. QxR
13. QxQ+
14. and must lead to mate.
15. Pawn up Qxp,
16. Nd4
17. to block it
18. and then the something
19. there must be be a potential mate in there.
20. Em just spotted the Bc1 there on
21. and that screws up the Rf4 move.
22. Em to way of getting that in there, play something like Rxc3
23. KxR,
24. doesn't really get me very far.
25. Bishop back into play.
26. My pawn on h4 is attacked,
27. doubly attacked by queen and rook
28. so I might advance my queen pawn to g5.
29. what can white play in terms of getting me in trouble
30. with Rxa7
31. Kxa7
32. pc7
33. Na6,
34. should be paying probably more attention to my k-side defence rather than my q-side defence.
35. Em very double edged position here.
36. Looking at bishop moves.
37. My Bf8 has to be a way of taking advantage of the fact that white has no squares for his king.
38. Eh, rook move
39. Rd3
40. can I hit the queen can I move the bishop?

S13 Position 4: W

1. Em, I looked at Nxf3 immediately
2. but that's no good because that leads to mate (Qxf3++).
3. I see black's king is in a spot of bother
4. and if I can play h4
5. with h5 protected
6. its mate.
7. I'm looking at a way to em, move my king from the square
8. because he's precariously placed at the minute as well.
9. I'm thinking about Qe1
10. but I think Bc5
11. defends rook adequately,
12. I think black has the possibility of creating a passed d-pawn
13. 'cause em, the pawn is quite advanced.
14. Apply more pressure
15. and I see a threat is Bb4
16. with a skewer with the rook through d-pawn.
17. So , something I'd have to be aware of.
18. Em, so now I'm looking for a counter to that move.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

19. But perhaps I play Kb1
20. if Bb4
21. Rxd3
22. eh I think Rxd3…
23. perhaps em, I think I well ok 'cause the Qxa5
24. and I think my d2 pawn's a gonner.
25. So I don't think that leading to any particular advantage.
26. Em consider playing Bd7
27. just to defend my c-pawn
28. and the bishop
29. and consequently the rook in the long-term
30. but that's not give me any threats
31. and its also easy to play Bb4.
32. I'd look briefly at Ne2
33. but that just gets taken.
34. I can't see a good way of removing the rook..
35. especially as its on dark square
36. as I don't have a dark squared bishop.
37. I think the Rf2
38. my pawn on as is attacked twice at the minute.
39. Em I'm thinking about sacking my rook on f6.
40. I'm, also bit doubtful how that works.
41. Also possible to play Rc5
42. but its only briefly
43. because its covered by the bishop.
44. Em, I don't see any checks.
45. I like the idea of playing Qb1
46. Bd4
47. Rxb1
48. Qxa5 '
49. cause now I realise I can play Rd5
50. no I can't.
51. trying to get my queen onto the light squares
52. maybe thinking about checkmate eventually
53. on f5 with my queen there.
54. Now I'm trying to think of a way to do that now. Em, perhaps I could…

S17 Position 5: B

1. (laughing). Complicated.
2. Count pieces 1-2-3-4-5-6, 1-2-3-4-5-6.
3. Actually kind of initially like certain aspects of my position here.
4. My king is very safe.
5. Em white has a lot of weak pawns eh on the queenside
6. whereas my pawns are safe.
7. I'd be wondering what happened earlier on in the game to get tripled e-pawns.
8. But eh and actually I'm not too happy with my Na7.
9. ok so ways for me to improve my position here.
10. Em well I'd like to get my Bg6 into the game,
11. my Na7 into the game.
12. Getting the Na7 sorry Na8 into the fame,
13. the only way to do that is via c7
14. and there are particular problems with that
15. because if I play Qb7
16. which is the way to do it
17. eh then eh white plays the move that white wants to play which is Rb3.
18. So I'd like to discourage Rb3
19. or how would I do it

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

20. so not eh Bh7
21. is a potential move
22. because that rook hasn't actually
23. Bh7 yea.
24. I kind of like that idea
25. because the white king is vulnerable
26. although white pieces around it n' stuff
27. so Bh7
28. when followed by kg7
29. followed by Rh6
30. followed by Qg6
31. followed by Qg5
32. looks extremely strong.
33. Actually that could be just winning.
34. Bh7 rook somewhere along the back rank.
35. Is there a particular problem with that?
36. There's certainly a problem.
37. Ah there's a problem with the whole idea
38. because eh pawn e5 will be un pres when I carry out that manoeuvre.
39. I'd have to look at precisely where the rook actually goes after Bh7
40. em it'd go to d8
41. because all the other squares whether allow me to,
42. are covered by
43. I gain tempos em now
44. I'd probably just pursue.
45. Oh the king can go to f5.
46. Perhaps a nicer way of doing it.
47. Potentially nicer anyway,
48. and black ok em wow there in this position.
49. I think Bh7 would definitely by my move.
50. 100% of the time
51. and after Rd8
52. I would definitely play either Kf5
53. or Kf(g)7.
54. Kf5
55. might have the …to allow Rd5
56. and which could give some kind of counter play against the d5 pawn.

S4 Position 5: W

1. Gosh, what is this?
2. Lets something…its completely random.
3. Phew, I'm white so I guess I should make a move.
4. My, all my pieces are weird.
5. The only good thing I can see is my e6 pawn.
6. I think I can queen that pawn.
7. I don't think there are particularly dangerous threats which black has, now …
8. one idea is to, a central idea is Rxf7
9. K x followed by
10. Kxc7
11. get Queen up,
12. eh probably wouldn't work at the moment,
13. now the other thing I might consider is Rf7
14. then with the same idea of playing e7.
15. Em, Rf7
16. Rx and
17. pxf7 my pawn is blocked,
18. but the good thing I might take the pawn on f5

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

19. (Qxf5).
20. Still my Re7 is probably a good one so I may not do it.
21. Now if I can move my pieces somewhere closer that should help
22. but I don't think I can move them,
23. ok so I should really find something…now so.
24. The material is level that's a good thing.
25. Oh well all my pieces should be probably better,
26. I have equally bad pieces
27. but I have at least a passed pe6.
28. em, centre Re8 and then
29. e7 perhaps,
30. now that's that's it. Am I finished? I am finished now phew.

S8 Position 5: B

1. Hmm, really wonder what this test is about? (laughs).
2. Ok an unusual position
3. and, more sensible than the last one,
4. black to move
5. although black looks like he's got a good pawn.
6. King is in a cage
7. but doesn't seem to be in danger immediately.
8. Pawn d3 looks very promising.
9. How do you get at it that is the question.
10. Ok, eh what's the plan?
11. Ok white, hmm ok Qb7
12. could be an idea try to get in on the second rank
13. and once the rook has moved
14. just play d2
15. followed by d1!
16. Does white have any threats?
17. Not really.
18. Pawns are weak, not going anywhere.
19. Ok the Ne/a8 not doing anything either.
20. Qb7 might be sensible
21. any way to get at the king,
22. h-file yea but that takes a long time.
23. Get rook out,
24. bishop out Kf5
25. looks quite safe.
26. Could get the king maybe,
27. the queen to h-file
28. and then try to mate with Qh4
29. but there are too many pieces in the way.
30. A more prophylactic approach is probably to play Qb7
31. Qb1/b2
32. then Rd1
33. just to get rid of knight on c1
34. and then queen the pawn.
35. Alternatively Kf5
36. followed by Qb1
37. d1
38. Qg4
39. could also be a good idea.
40. em yea, yea Qb7 looks like the move that I would play in this position.

S1 Position 5: W

1. Ah, ok here we have another extremely unclear position
2. which is actually possible to occur….
3. Well likely… well anyway the actual position as I see it.
4. Once again the kings are in very compromised positions.
5. Black is more active than white,
6. white seems to have more immediate position plusses.
7. The Rd7-e7 em
8. what white needs to do there is to somehow… untangle his forces on the kingside seems to be maybe to do…
9. the actual opportunity where for instance where eh… moves such as Kc4
10. followed by c3
11. is not immediately observable
12. should actually… perhaps with the idea of b5
13. followed by Kb4
14. em would be interesting
15. as em the R +
16. the queen then would come into the play on the third rank.
17. And while the king is on b4
18. c4 would be possible
19. followed by c5+
20. em king forced a7
21. and after a move like Rd8…
22. em mate
23. or check on the a-file
24. it would be mate…
25. even b6 +
26. px
27. px
28. Kx eh
29. Rc3
30. followed by Rc6+
31. sacrificing rook
32. bring the queen in in some lines,
33. but that being said… instance Kb5
34. c4
35. c5+
36. Rxc7
37. right were black as black do anything to counteract this?
38. Ideally, if playing b5
39. black could…
40. why I like the idea of playing b5
41. is black had looked to take advantage of the weakness on the dark squares…
42. by just opened up such as eh eh by playing Qa8 to a5+
43. with the queen misplaced on the other side of the board,
44. and it doesn't look like its going anywhere fast…
45. but there is the threat of the h-pawn
46. em but perhaps key to blacks defence of this position is actually after b5
47. to play a move like Ka5
48. to stopping the idea of Ka4
49. it looks quite absurd…
50. but maybe it is not as stupid it looks…
51. unless if white was to take on b5…
52. then Rxc7
53. would surely give a big advantage to white
54. followed by Rc4…
55. Rb4+
56. and just eh…
57. Bc4+ clearly be,

58. the main problem that white has is in this position the Rb1
59. while shuts off the king on b3
60. and divides the board in two.
61. As the solid pawn structure is defended all of the action is taking place on the a: B and g files.

S2 Position 5: W

1. Right, … well again I see a position where white has an advanced passed pawn
2. and this time its important.
3. I'd probably would be thinking if it doesn't win the game what black will have to give up to stop it queening,
4. so, … the first move I'm thinking of is Rf7
5. or Rd7.
6. Rd7 looks better
7. because if I go to Rf7
8. RxR
9. pxR
10. forces knight to go away,
11. but at the same time the queening square is covered by the queen.
12. Rd7
13. is a little more threatening
14. especially if I can get… something else behind it
15. like the other rook
16. which at the moment I can't
17. because of the eh p f4.
18. So, that would be the lines I'm thinking of ,
19. em what threats has black got against me?
20. Eh, seem to be pretty well caged in by all my own pieces
21. and look safe
22. unless I'm misunders' estimating.
23. One move he has maybe is Bxp+
24. drawing me out
25. but I don't see where it leads him.
26. So , I would be thinking of Rd7
27. protecting a further advance of the pawn (e7)
28. dividing the game
29. I would try to angle myself back somewhere to Qe2
30. to protect it further,
31. but I know I couldn't stay there long
32. because he could just push the pawn up against it
33. and eh force me to move
34. but I would be looking along those lines.
35. Maybe not getting in to get a queen
36. but of forcing black to give up a piece
37. soon maybe eventually end of the game win on material…
38. that's em, is that 3 minutes?
39. (E: no, you can keep going if you want).
40. The black rook looks dangerous
41. but I think its just looks,
42. em because even if I have to move my queen at any stage
43. my bishop is protected by my rook and queen.
44. And I can see the knight protecting the pawn.

S20 Position 5: B

1. Ok, I take it I'm black again here,
2. and again very wild open position

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

3. with the kings out of the way.
4. Now, I have to say for starters.
5. I'd see black with three pawns on are eh file
6. so, you can look at that and say is that going to go anywhere,
7. whereas, actually has two passed pawns.
8. Now they've a bit to go yet
9. and I'd say that this is going to be difficult maybe to stop,
10. eh also the Rg8 is certainly pinning the king in there
11. if there was to be any sort of.
12. Where white will go is difficult to say,
13. there is his, eh Bb8
14. the white bishop is probably not that well placed there
15. and I think the the eh… what would he play,
16. his Na8 is trapped.
17. So really he is going to have to try to extricate that Na8,
18. the queen is defending pa5…
19. oh no sorry that pawn is defended by the queen
20. if he moves back there's the possibility of moving.
21. Eh, I think… again the prospects look like it has to be also
22. the black Rd2 has no support
23. and no defence
24. and I think the queen will have to become more active
25. and still its going to let that pawn advance
26. so what can we do to stop that…
27. the thing is the rook the black rook really has nowhere to go.
28. Move onto the back rank.
29. That might be the best thing,
30. all things considered
31. eh with the possibility on the h-file
32. and still is not threatening or doing anything.

S11 Position 5: B

1. Ok, the first thing that strikes me looking at this position is that it certainly didn't occur in a real game.
2. It's a composed position.
3. My first impression as I look at the board is that from the wrong side?
4. So its very difficult.
5. Looking very complicated.
6. Very hard to think about the position.
7. I see that em, that my king is safe,
8. maybe not as safe as white's.
9. Ah actually my king may not be safe
10. because its if white can somehow generate threats against it somehow,
11. it cannot go back to g7,
12. em also eh white has eh maybe has some threats of advancing his pf4
13. and eh queening eh hmm.
14. I em, don't see any particular, any way of getting eh at white's king.
15. I don't see anyway for white to get at mine,
16. eh I I my knight h1 (a8) stuck in corner.
17. I can't see how I can move that.
18. I could possibly try to re-manoeuvre my queen
19. by playing Qb7…
20. eh now with possibility of coming down to g2 (b2).
21. Well that would probably force off an exchange of queens,
22. em but why do I have to be in so such a hurry,
23. I could play Qg7 (Qb7)
24. with threat of going to b2
25. but not immediately.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

26. I can then re- deploy my Re6 to em h6…

S12 Position 5: W

1. Ok, very complicated position here.
2. Lots of pawns 1-2-3-4-5-6, white six pawns.
3. White Nh1 Qh3.
4. Rook stuck up there on e7.
5. Eh black has Rb1.
6. King exposed although maybe not that bad.
7. This position here does not look good.
8. Black has pawns on f-file.
9. Ok they're doubled
10. but they're eh fairly clean run through to promotion.
11. Same could be said about my pe6
12. eh what how can I advance that.
13. Rook supporting it.
14. Immediate danger is from my perspective bishop moves
15. Bxd4+,
16. … Rxe6
17. pxN,
18. black isn't going to play that,
19. bishop into play
20. Rook Bishop out to get his bishop active
21. drop that…
22. potential moves for white
23. Rf7
24. RxR
25. pxR
26. no and queen just wins that back
27. em, Rf3 attack pawn
28. what against it can be done.
29. Rf3
30. Nf2
31. knight active,
32. Ra1
33. Bxd4
34. RxB
35. back pawn up pf3
36. white's Nf8 is fairly inactive…
37. Ng6 to hit my rook.
38. Where does my rook got then, Rxc7…

S13 Position 5: W

1. Em, ok I see I've got tripled pawns on the d-file.
2. I see I've got a protected e-pawn on e6.
3. I would consider playing Rxf2 (c7)
4. I'm also considering playing Ba4
5. at some stage.
6. Ba4 em,
7. … just think that's a good square for the bishop.
8. I'm thinking about playing Ba2 em,
9. perhaps that's a little pointless.
10. I think about playing Kc4
11. with the idea of Rc3
12. and then Kd3.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

13. I'm also thinking about trying to vacate the d3 square for the queen.
14. So maybe I'll play Rf3
15. em, with the idea of Qf1
16. attacking-
17. ok I can't play that.
18. Rf3
19. with the idea of Rf1
20. Rxf1
21. Qxf1
22. Bh2
23. Qc4
24. maybe something like that
25. with a strong attack on c2, c7 pawn.
26. Em, I', thinking of ways to support advancing the e-pawn,
27. em I see my knight is dominated by the bishop
28. but the black bishop also relatively out of play.
29. Em my king position is a little restricted.
30. I'm thinking about transferring a rook somehow to the a-file
31. keeping blacks king locked in,
32. so I'm thinking about playing Kc4
33. followed by Ba2
34. rook moves
35. and then Ra3
36. with the idea of following all that up with Qd3
37. and aim for mate
38. or something like that.
39. The black king is very restricted.
40. I don't see any threats for black immediately
41. and looking for them,
42. eh I see he's double pawns on c-file
43. but they are passed pawns.
44. I control the queening square
45. and the foremost pawn is well supported.
46. Em I'm thinking about playing Qg2
47. attacking the bishop.

S17 Position 6: B

1. Again I'll count pieces 1-2-3-4-5, 1-2-3-4-5 (counting pawns).
2. We're even on material
3. em and its my move
4. and I've to work out whether I can take that R on h6.
5. Basically because if I take that rook its game over
6. (N.B. the rook was half on h6/h5, I then adjusted it and chess clock).
7. Em, ah (E: sorry about that, S: laughing).
8. Ok, eh ok we have an interesting position then,
9. my kingside is a bit vulnerable
10. and white's king does seem to be safer.
11. My Bf8 isn't a good piece
12. because, can't really isn't getting into the game at the moment
13. em but I do have positives in so far as Ba6.
14. White's Ba6 isn't great
15. and em I've a nice passed pd4.
16. So what kind of things would I want.
17. Well I think an endgame would suit me in this position
18. because em, the pd6 would then become a weakness
19. and my king would be less vulnerable.
20. So do I want to look at playing moves like Qe1

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

21. which is position ally will swap queens off
22. because the knight has to move.
23. Em the knight doesn't have many good squares.
24. Oh then the d pawn falls
25. so that might force white into doing some radical action.
26. Ok Qe1
27. em what kind of radical action does ,
28. oh sorry my knight would be unpres from my queen.
29. Ok, how do I sort that one out?
30. Em d6 looks just equal.
31. He just,
32. I can take the eh Ba6 with the knight.
33. It gets rid of my knight.
34. That's one way of want to play.
35. I could play,
36. play Rb6 (8?)
37. attacking the d-pawn again.
38. The interesting problem with that is I would be worried about a few moves for white in reply to Rb6.
39. e5 which protects the pawn
40. and threatens to open up my kingside
41. and gives a good square for the Nf2.
42. So ok Nxa6 looks like it could be a move then.
43. Rxa6
44. Rxa6
45. Qe1,
46. oh see that allows Nf2 going to a difficult square…
47. d3
48. em there's actually go to be some tactical difficulties
49. because here like Rc7
50. is a threat for white.
51. Maybe I should just move the Ne6
52. stops opening up my king,
53. em, problems.
54. Ne6
55. Na6 wins the pawn
56. allowing me to build up slowly on the d6 pawn,
57. push a pawn.
58. c7 is a problem there.
59. Ne6
60. white can play e5 in that position.
61. I think that's opening the white king.
62. I'd be better with the knight over protecting my king,
63. better, ok don't need the extra time
64. Ne6 is the move I would play.

S8 Position 6: B

1. Eh, ok again quite messy.
2. Why aren't I getting any positions that I like?
3. Controlled positions with clear simple plans.
4. Ok so what does this position look like?
5. Ok d7 is covered.
6. Material equal,
7. d-pawn looks strong,
8. b-pawn not so,
9. em the end position looks,
10. yea quite good for black,
11. em the only problem is d7

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

12. we have to take care of that eh otherwise,
13. the e5 square would be very nice for the knight
14. but he can't get there.
15. Em: Black's king looks quite safe.
16. Ok the obvious move would be d3
17. and say, now does that have any disadvantage?
18. Ok, eh could xc5
19. xd7,
20. that doesn't really do anything.
21. He can always hide there on g8,
22. hmm ok so still have a-pawn.
23. d3 looks like obvious move.
24. I wonder what Michelle is thinking there? (laughs).
25. Ok tactical solution here
26. xf2,
27. Qe1 is also possible.
28. Qe1 getting rid of that Q first
29. could make the whole thing safer even?
30. Yea how about Qe1,
31. what happens then?
32. Qe1
33. where does the knight go.
34. Knight doesn't go anywhere,
35. go to h3.
36. Its probably better even yea
37. maybe play Qf1/h1
38. followed by Rg2.
39. Kf5
40. is that really dangerous?
41. Qg6+
42. oh perpetual check is coming up.
43. Yea and I see so we do have to be careful.
44. Ok Qe1
45. Kf5 ok
46. if we take Qg6+,
47. now that looks too dangerous,
48. at least perpetual.
49. Ok the, what do we do now, have to hurry up.
50. Have to find something against Kf5
51. so maybe you need a move like Qc4,
52. yep I probably Qc4,
53. Rd5 blocks that!
54. Hmm not so simple at all!
55. Qg5
56. Qg6+,
57. quite tricky.

S1 Position 6: W

1. Now eh once again its very unclear position.
2. White's queen and rook are quite a way far down the board…
3. g6 is threatened queen
4. gaining an essential distilled advantage.
5. There seems to be some danger though for white with this position
6. at the that eh black is fairly well placed to launch some kind of eh desperate attack. The Rh2 on the seventh em looking at e2
7. which is the only…the only only pawn separating the rook form the king.
8. A move like Kc4

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

9. for black with threaten
10. Qb3 for black
11. then Kb1,
12. Kb1
13. followed by eh Nb5
14. which immediate mate
15. while the Na3+
16. Ke1
17. Na3+
18. Ra1++
19. a nice little combination.
20. I'm not sure what white can actually do against eh Kc4 threat
21. but I'm thinking about b3
22. to attack the rook
23. and in Ra3+
24. Bb2
25. and perhaps a move like a4 anyway…
26. just probing for more reasons
27. a4
28. followed by a3
29. and Rxb2+
30. and pxa4
31. might be followed by Kc4 again.
32. Em this time with the Qxb3+
33. coming probably resulting in mate.
34. So white probably has eh a couple of problems here
35. although it looks like say well…
36. the weakness of the light squares is is the key in this position.
37. Another thing I'm quite tempted to sacrifice, ach to consider sacrificing eh a rook for knight
38. Rxc7
39. Qx
40. followed by Ne6
41. Nd4+
42. eh of course it won't work
43. because Rxe2+.
44. So probably the best idea in this position is for white to try and take control of the light squares
45. and he might be be able to play moves such as Qf5
46. Qb3+
47. at which point there will be possibility of of eh playing em playing the knight back into play.
48. The only drawback with that eh is that the Ra4 is hitting the Nf4
49. white now has to be very careful as to how he does it…
50. course the other thing that would be possible Qf5
51. d3
52. followed by NxB
53. and if Rx Ba4
54. give white eh clear advantage
55. as the e3 pawn is quite weak
56. em although his bishop is poorly placed on c1…
57. the bishop is actually doing a good job of protecting…

S2 Position 6: W

1. Ok, now this position… I see as very even again.
2. Eh, I'm looking at the business of the knight attacking the pawn in the centre
3. and the other pawn being pinned in a way against the queen protected by a king.
4. Also protected by the knight.
5. So the lines I'm looking at at the moment is to
6. RxN…

7. it's a dramatic type of move
8. but that doesn't give me much against the king.
9. Now we need to have something focussed on… the queen…
10. so that would give me, if I prepare to shift the queen off the square
11. Bxp on e3
12. moves Qe3
13. if he takes
14. and I'm not sure whether that gives me what I'm looking for…
15. eh black is dangerously close here
16. with the pawns and also, …
17. eh has a rook bearing right down on my king position…
18. so to make a move… I would … leave the pe3
19. and I would try to advance the e-pawn myself,
20. because I have lots of squares covered all the way down to the queening square
21. bar the last one
22. and if I was to kind of take off the knight at an opportune stage
23. it would be looking very good for me.
24. So, I probably should analyse the black position a bit more in terms of the threats that…

S20 Position 6: B

1. Once again I'm black and the assessment is that black has 2-5 pawns two minors (minor pieces) two rooks and queen.
2. White has queen, two rooks, two minors 2-4-5 pawns.
3. So material looks pretty well equal.
4. And black effectively has two passed pawns
5. if he can do anything with them.
6. So I think that is a possibility now
7. so far white's attacks are concerned he hasn't anything very much.
8. He has the queen
9. which isn't going a whole lot of places…
10. eh the knight black Nc5 has prospects
11. and might be worth bringing it into play.
12. e3, e6 sorry.
13. Eh it doesn't effect anything immediately
14. whereas at the moment if the pawn advances it is supported
15. but its already well supported with the eh, … yea…
16. oh the trouble is if that pawn advances on d3
17. white might well then play eh his pf3 to f4
18. which means the pawn would be pinned against the queen
19. and eh, I don't eh yea.
20. White can't very well move the queen
21. because the knight is threatened then by the rook
22. and let me think.
23. All things considered
24. I'd be inclined to play d3, with the pawn.

S11 Position 6: B

1. Well on first glance it looks like my king is in some trouble.
2. My position black has some threats
3. or white has some threats along the h-file
4. and also along the seventh rank for example
5. he would like to play, I think eh
6. Rxe7
7. only for my knight is protecting that.

8. So em on the other hand I have some pressure against em against white position, white's king is not in a favourable position.
9. He seems to be leading the attack with the king
10. so, that might be of some benefit to me.
11. Em I think eh, a move I would look at, em I would actually like to play
12. Ne6.
13. Once again I have this problem with the eh e-pawn.
14. Em, but em, I think I could probably play this move em, (E; keep on talking),
15. eh I think I could probably play this move.
16. I could actually, there is another move I could play
17. Nxa6
18. and recapture the b-pawn
19. which might be an idea.
20. Actually I couldn't capture the b-pawn
21. because Ra5 cover it.
22. Em, possibly a good move might be to play Qd2
23. attacking knight
24. threatening to win the Nf2.
25. It is defended by the queen
26. and I don't see any other way for em, for white to hold on to Nf2.
27. He can play, he can play
28. Nh1,
29. eh I would also like possibly,
30. to bring my bishop into the attack if I could,
31. eh by playing Bxd6
32. except for that the queen is covering that,
33. so I think, I think a good move might be Qd2,
34. em I don't see any threats
35. any immediate
36. threats for white is Rh5,
37. so probably go after that.
38. I think, I think that would probably be best.
39. I think the position seems to be safe enough.
40. Qxd2 em, em bishop…

S12 Position 6: W

1. Ok, I have , I'm white I have 1- 5 pawns, against 1-5 pawns.
2. Eh I have bishop and knight vs bishop and knight two rooks queen.
3. Opposite colour bishop.
4. Queen quite advanced
5. Rg7 quite advanced
6. but black has Rh2
7. also well advanced on the seventh rank.
8. Black could play Rxe2+
9. But that's protected by the knight.
10. Can I attack in perspective, could play eh,
11. Nxd5
12. NxN
13. not good for me.
14. Pushing my pawn on my e5 pawn come racing through
15. so if I, can white sack the exchange
16. so white push pain.
17. RxN
18. QxN
19. Rxe2+
20. takes away the knight from my protected e2 pawn.
21. Nothing in that for black.

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

22. Ok, Kb1 maybe reasonably well protected
23. Q attacks on b2,
24. so there's a danger of that happening
25. em, yet I need to be aware of that
26. so Rh2 attacking e2.
27. R+
28. give me big big problem.
29. But the Bc1 protects d2
30. so swing knight into action.
31. e6 looks like a reasonably good move.
32. Eh Re4
33. Queen back push on again,
34. still can't take,
35. push again Ne8
36. do something about that.
37. Re4 looks like a reasonable move.
38. In this position I think I would play e6. yea…

S13 Position 6: W

1. …Em ok so
2. NxB
3. px
4. RxB
5. and I see NxB
6. Ne8
7. a strong intermediate move threatening the queen,
8. Qc5
9. ok queen ok
10. so I'm now considering Qg7
11. defending the eighth square
12. stopping the fork
13. and threatening Qd3+
14. em, I'm also looking at the possibility after Qd6/f5 of
15. playing Kg3
16. so I'm attacking the rook
17. and if I play Qd3+
18. the king moves or approaches the rook.
19. So Qc5
20. Kh6
21. Qd3+
22. Rc4 (E: is that Qb3+ or)
23. No Qd3
24. Qf5..
25. R…
26. ok I'm looking at anyway that black can destroy get rid of my Nf4
27. which is defending my vital pawn e2.
28. so perhaps he can play…

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Appendix D: A sample of twelve problem behavior graphs (three masters and three novices for the three normal positions and the three random positions).
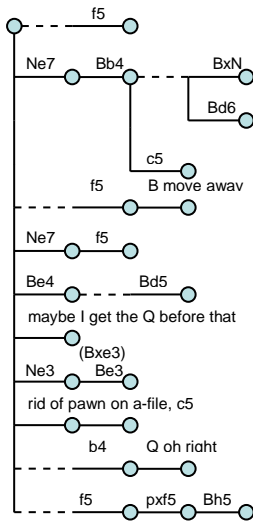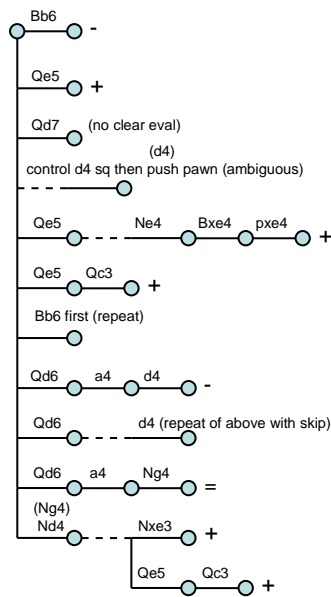

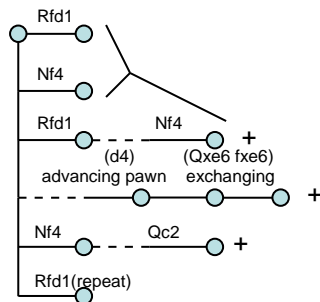Example 1: Grandmaster (participant 4), normal position 1, black
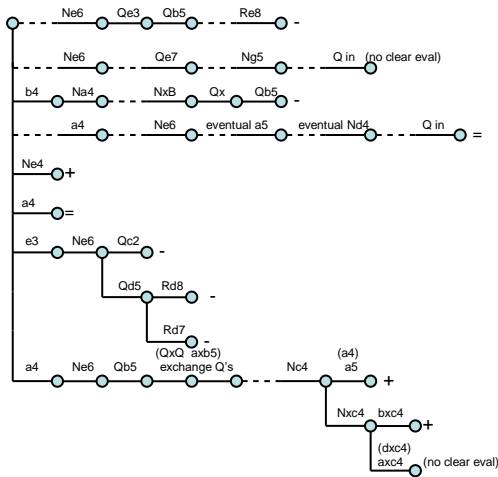
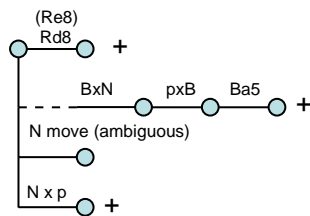Example 2: International Master (participant 1), normal position 1, white

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Bb6 –

Qe5 +

Qd7 (no clear eval)

(d4)
control d4 sq then push pawn (ambiguous)

Qe5 Ne4 Bxe4 pxe4 +

Qe5 Qc3 +

Bb6 first (repeat)

Qd6 a4 d4 –

Qd6 d4 (repeat of above with skip)

Qd6 a4 Ng4 =

(Ng4)
Nd4 Nxe3 +

Qe5 Qc3 +

Example 3: Fide Master (participant 17), normal position 2, black

Rfd1

Nf4

Rfd1 Nf4 +

(d4) (Qxe6 fxe6)
advancing pawn exchanging +

Nf4 Qc2 +

Rfd1(repeat)

Example 4: Novice (participant 11), normal position 2, white

(Re8)
Rd8 +

BxN pxB Ba5 +

N move (ambiguous)

N x p +

Example 5: Novice (participant 12), normal position 1, black

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.
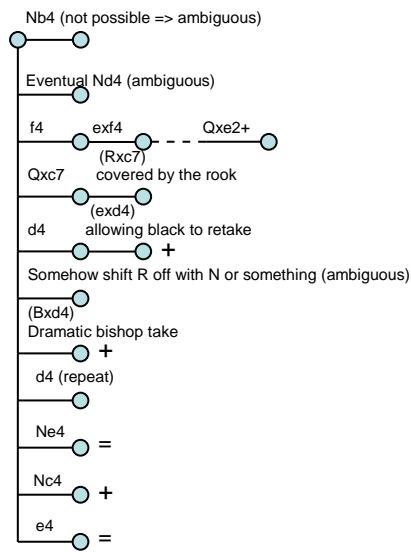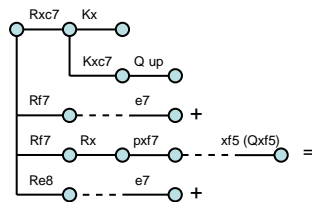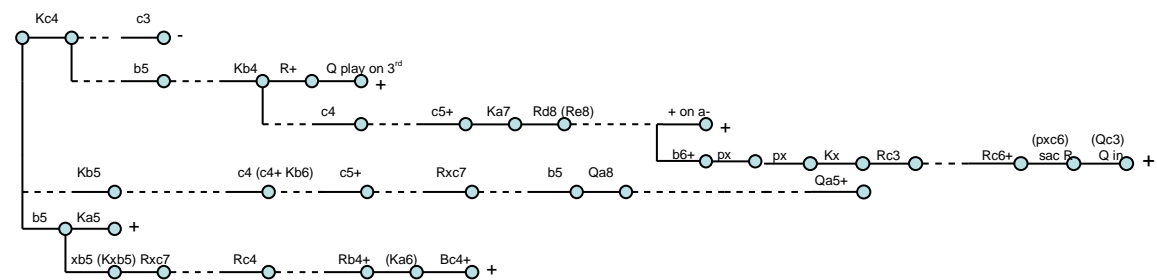
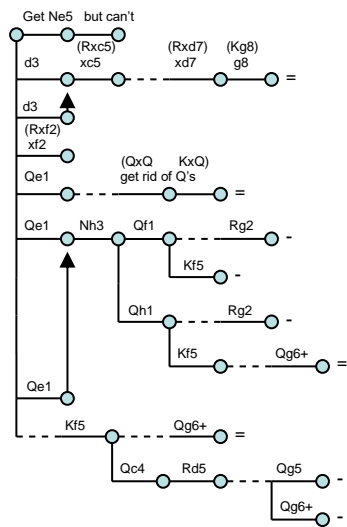Example 6: Novice (participant 2), normal position 3, white

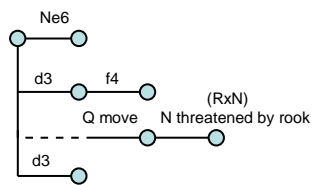Example 7: Grandmaster (participant 4), random position 5, white

Example 8: International master (participant 1), random position 5, white

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Example 9: International master (participant 8), random position 6, black

Example 10: Novice (participant 20), random position 6, black

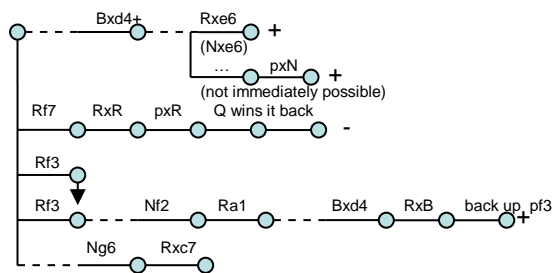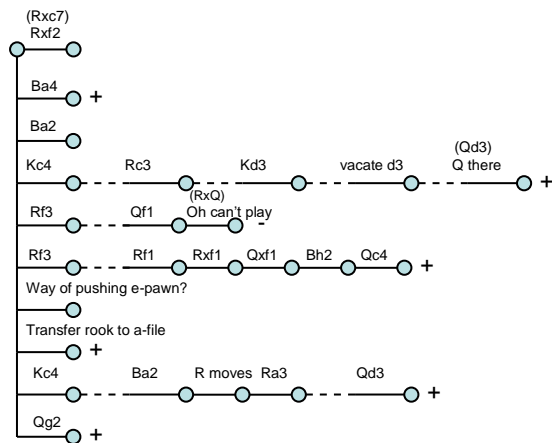Example 11: Novice (participant 12), random position 5, white

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

(Rxc7) Rxf2

Ba4 +

Ba2

Kc4 — Rc3 — Kd3 — vacate d3 — (Qd3) Q there +

Rf3 — Qf1 (RxQ) Oh can't play -

Rf3 — Rf1 — Rxf1 — Qxf1 — Bh2 — Qc4 +

Way of pushing e-pawn?

Transfer rook to a-file +

Kc4 — Ba2 — R moves — Ra3 — Qd3 +

Qg2 +

Example 12: Novice (participant 13), random position 5, white

Appendix E: The experimenter's think-aloud script

Instructions
Experimenter:

"Thank you for participating in this study. I would like to remind you that your participation is anonymous and the data will be handled strictly confidentially. Should you wish to withdraw at any stage you are free to do so. This study is interested in the way chess players think about chess positions, and not in unconscious emotions or hidden thoughts.

You will be shown 7 chess positions including one practice position to ease you into the procedure. You will have 3 minutes to think out loud about each chess position. If you are finished thinking about the chess position before the time is up please feel free to say so. I will be recording you with a dictaphone and taking notes while you are thinking aloud. When your time is up I will ask you some questions about what you were thinking in the time.

In a moment you will be shown a (practice) chess position. You are asked to choose a move you would play in the way you are used to going about choosing a move in a real game. When time is up you will be asked to declare your chosen move. It is important that you say aloud everything, everything…, that you think while choosing a move. If you stop talking at any stage I will prompt you with the words 'keep on talking'. Also there is no need to explain why you are thinking about something, it is more important that your protocol is natural than comprehensible to me. Would you like to ask me any questions at this stage?"

Response:

Experimenter:
"Are you ready to start?"

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Stoppage:

After a 3 second stoppage the first one to two seconds could be "ok, what do I do now?" on the third second the participant is likely to start doing something or looking for something so the words "keep on talking" was uttered quietly as outlined in van Someren et al (1994).

Retrospective Evaluation Script

"Now I will ask you some questions about what you were thinking in the time you were analysing this chess position. I will remind you of some of the moves/move [perhaps a subject particularly non-expert will just analyse variations for one chosen move] you considered. I'll ask you to tell me whether or not a particular move led to a positive outcome for your position, in other words, an improvement of your position from its present state, or to a negative outcome for your position, in other words, to a worsening of your position from its present state, or whether it neither positively nor negatively affected your position. Is there anything you would now like to ask me about that?"

Response:

Experimenter:
"Okay, your first move was: (**x**)
"Moves that followed on from (x) were (y) and then (z)
                                        (r) and then (s)
                                                ...''
"Did the line {**x**, y, z, ...} lead to a positive or negative outcome, or neither?" (For example did the line **Ne4**, Bf5, Nd2 lead to a positive or negative outcome or neither?)

Response:

Experimenter:
"Did the line {**x**, r, s} lead to a positive or negative outcome, or neither?" (i.e. **Ne4**, Bd6, Qd5).

Response:

Single moves were also indicated.

Experimenter:
"So, did the move (**x**) lead to a positive or negative outcome, or neither?"
(e.g. "So, did the move Ne4 lead to a positive or negative outcome, or neither?")

Response:

Initial instructions for the remainder six positions were then modified as follows:

Experimenter:
"Once again I will ask you to think-aloud while choosing a move in a chess position. Again it is very important that you say everything you think while thinking aloud. If you are finished thinking about the position before the three minutes are up, feel free to let me know. Ok are you ready to start?"

Cowley, M. (2016). Chess masters' hypothesis testing in games of dynamic equilibrium. *Econometrics: Econometric & Statistical Methods – General eJournal*, Jan 13, 2016. See also *Game Theory & Bargaining Theory eJournal*, Jan 13, 2016.

Response:

Experimenter:
"Again I will be reminding you of moves you looked at while choosing a move in this position and asking you to tell me whether they led to a positive negative or neither positive or negative outcome for your position. Ok are you ready to start?"

Response:

Verbal Debriefing

"Thank you once again for participating in this study. Now I will tell you about what the study aimed to examine. In everyday thinking people generate hypotheses or ideas about incidences and relationships in the world around them. Psychologists have found that people tend to look for evidence to confirm their own ideas rather than look for evidence to prove their ideas false. This study was interested in how different levels of expertise affected confirmation or falsification of ideas in a specific domain — that is, the domain of chess. Is there anything else you would now like to ask me?"

Response:

Appendix F: The experimenter's recording sheet

Chess Position Number: _____   White/Black side: _____

Move chosen overall: _____

Evaluation box ☐
+ : positive outcome for subject's position
- : negative outcome for subject's position
= : neither positive or negative outcome for subject's position

Move/s chosen	Variation/s examined		Variation Evaluation Explanation +/-	Overall Move Explanation +/-