Brain and Cognition 80 (2012) 160–169

Contents lists available at SciVerse ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Dissociable stages of problem solving (I): Temporal characteristics revealed by eye-movement analyses

Kai Nitschke [a,b,1], Nina Ruh [a,b,1], Sonja Kappler [a], Christoph Stahl [c,d], Christoph P. Kaller [a,b,*]

[a] Department of Neurology, University Medical Center, Freiburg, Germany
[b] Freiburg Brain Imaging Center, University of Freiburg, Germany
[c] Department of Psychology, University of Cologne, Germany
[d] Department of Psychology, University of Freiburg, Germany

ARTICLE INFO

Article history:
Available online 20 June 2012

Keywords:
Eye movements
Problem solving
Planning
Tower of London task
Prefrontal cortex

ABSTRACT

Understanding the functional neuroanatomy of planning and problem solving may substantially benefit from better insight into the chronology of the cognitive processes involved. Based on the assumption that regularities in cognitive processing are reflected in overtly observable eye-movement patterns, here we recorded eye movements while participants worked on Tower of London (TOL) problems that comprised an experimental manipulation of different task demands.

Single-trial saccade-locked analyses revealed that higher demands on forming an internal problem representation were associated with an increased number of gaze alternations between start state and goal state, but did not show any effect on the durations of these inspections of the states. In contrast, higher demands on actual planning in terms of mental manipulations of working memory contents coincided with a prolonged duration of the very last inspection of the start state (i.e., immediately preceding movement execution) but did not show any effect on the number of gaze alterations.

Differential task demands on internalization and planning processes during problem solving hence selectively affect different eye-movement parameters. Moreover, these findings complement previous neuroimaging data on dissociable contributions of left and right dorsolateral prefrontal cortex in problem solving with novel evidence for a corresponding dissociation in the eye-movement patterns reflecting the associated cognitive processes.

1. Introduction

In its simplest form, problem-solving behavior may rely solely on trial and error. Yet, in many situations beyond everyday routine, the ability to plan ahead is crucial for effective performance as it allows us to bridge the gap between the present and the future. In particular, planning entails the mental modeling and anticipation of the consequences of actions prior to their execution in the real world (Goel & Grafman, 1995; Ward & Morris, 2005). It is well known that the dorsolateral prefrontal cortex (dlPFC) plays a key role in the cortical network involved in planning and problem solving (Owen, 2005). In a functional magnetic resonance imaging (fMRI) study on the Tower of London planning task, Kaller, Rahm, Spreer, Weiller, and Unterrainer (2011) have recently shown a double dissociation between the function of left and right dlPFC and two different cognitive aspects of problem solving: Stronger activation of the left dlPFC reflected higher demands on structurally analyzing propositional information by identifying the relevant differences between the start and the goal state. In contrast, stronger activation of the right dlPFC was associated with higher demands on mediating the integration of information into a coherent action sequence. In other words, left and right dlPFC were found to contribute differentially to presumably separable cognitive stages of (i) establishing a mental representation and (ii) actual planning processes, respectively.

In the fMRI study of Kaller, Rahm, Spreer, et al. (2011), these two different aspects of problem solving were implemented using experimental manipulations of two Tower of London (TOL) problem parameters, namely *Tower Configuration* (see Footnote [2]) and

---

* Corresponding author at: Department of Neurology, University Medical Center, Breisacher Str. 64, 79106 Freiburg, Germany. Fax: +49 761 270 54160.
E-mail address: christoph.kaller@uniklinik-freiburg.de (C.P. Kaller).

[1] These authors contributed equally.

[2] In the present manuscript, we use the term *Tower Configuration* instead of *Goal Hierarchy* that had been previously used by Kaller, Rahm, Spreer, et al. (2011). This is done in order to emphasize that, at least for simple three-move problems with an intermediate move, it is not the ambiguity of the sequence per se that leads to latency effects on initial thinking times (cf. Kaller et al., 2004) but rather its identifiability and the related perceptual difficulty of matching and integrating information between start and goal states. Consequently, performance and/or latency effects following experimental manipulations of *Goal Hierarchy* in easier and more difficult TOL problems may not be attributable to the same underlying cognitive operations (Kaller, Rahm, Köstering, et al., 2011). For further details, see also Section 2.

http://dx.doi.org/10.1016/j.bandc.2012.05.003

Box 1. Different aspects of problem solving in the Tower of London (TOL) task.

The Tower of London (TOL) is a so-called disc-transfer task in which start and goal states are represented by three differently colored balls placed on three rods. The aim of the task is to transform the given start state into the goal state within the minimum number of moves. Thereby, task demands in the TOL are determined mainly by two structural task parameters: *Tower Configuration* and *Search Depth* (Fig. B1, panel rows and columns, respectively; cf. Kaller, Unterrainer, Rahm, & Halsband, 2004; Kaller, Rahm, Speer et al., 2011; Kaller, Rahm, Köstering, & Unterrainer, 2011).

In the TOL, goal succession can be more or less ambiguously derived from the *Tower Configuration* of the goal state. If all balls in the goal state are stacked on one single rod (full goal tower, Fig. B1), sequences of placing the balls into their goal position can be deduced unambiguously as the bottommost ball has to be put there first, followed by the second and, finally, the topmost ball. If, instead, balls are distributed across more than one rod (partial goal tower, Fig. B1), goal sequences become partially ambiguous, requiring matching and comparing of locations and identities of balls across both start state and goal state in order to infer the appropriate succession of goal moves. Therefore, different *Tower Configurations* posit different demands on building-up a mental representation and internalizing the problem situation.

In contrast, *Search Depth* refers to the actual mental generation and evaluation of sequences. In the TOL, problems with an intermediate move imply mentally looking ahead, because the generation of individual moves is required while taking into account their interdependencies. As illustrated in Fig. B1 (left side, problems with intermediate move), neither the gray nor the white ball can be placed directly into their final goal position, but there are two equivalent alternatives for initially moving the gray ball in order to release the white. However, the interdependency between the gray ball and the white ball becomes obvious only by mentally looking ahead as only one of the two alternatives for initially moving the gray ball allows to solve the problem within the minimal number of moves, whereas the other will block the subsequent goal move of the white ball. In contrast, problems without intermediate move can be solved in a simple straightforward manner by placing each ball directly in its goal position without the necessity of mentally looking ahead (cf. Fig. B1, right side).

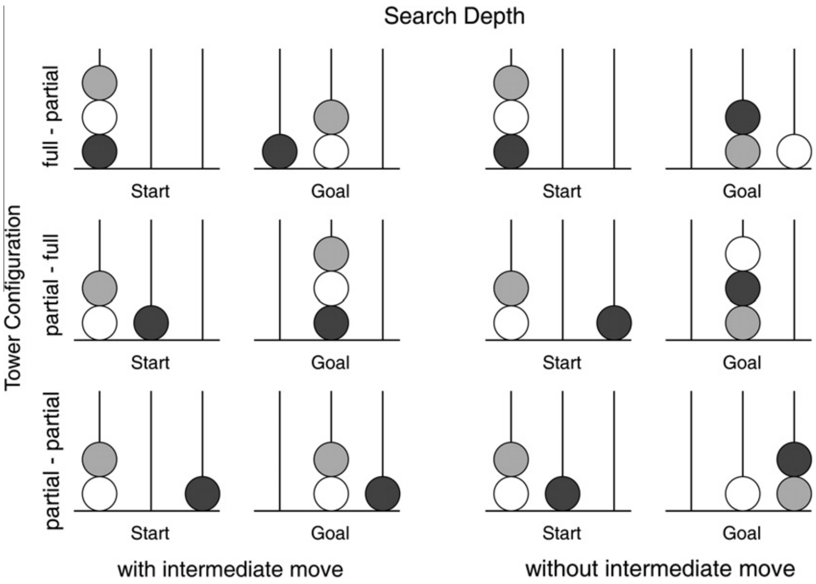

Fig. B1. Experimental design comprising a 2 × 3 factorial manipulation of the factors *Search Depth* (columns: with intermediate-move/without intermediate-move) and *Tower Configuration* (rows: full start–partial goal/partial start–full goal/partial start–partial goal). *Search Depth* involves dealing with intermediate moves and associated interdependencies between subsequent move alternatives. Intermediate moves occur if a ball cannot be placed directly in its final goal position, but has to be temporarily deposited on another rod. *Tower Configuration* defines the physical setup of start and goal states. Full towers contain three balls positioned on only one rod, whereas partial towers comprise an arrangement of three balls distributed on two different rods. See main text for more details.

Search Depth (for a comprehensive illustration of the two parameters, see Box 1). Based on the assumption that systematics in cognitive processing are to some extent reflected by consistencies in eye movements, previous research suggests that these neurally dissociable processes during problem solving do not completely proceed in parallel. More specifically, it is proposed that they may follow an at least partially sequential order, with internalization processes of constructing a mental problem presentation terminating before planning processes (Hodgson, Bajwa, Owen, & Kennard, 2000; Kaller, Rahm, Bolkenius, & Unterrainer, 2009). For instance, Kaller et al. (2009) demonstrated that actual planning (as elicited by higher demands on *Search Depth*; but see also Footnote [3]) was exclusively associated with longer durations of the last inspection of the start state before movement execution (none of the precedent inspections of the start state or the goal state revealed any significant relationship with planning at all). Based on these results, Kaller et al.

---

[3] In the present manuscript, the term *Search Depth* refers to the structural manipulation of “with” vs. “without intermediate move” in three-move problems as applied in Kaller et al. (2009).

(2009) suggested that the precedent inspections most likely reflect the formation of an internal representation of the two tower configurations including processes of comparing and matching individual ball positions in order to identify the relevant differences that state the actual problem. That is, *internalization* comprises both the establishment of a mnemonic representation of the problem configuration and also the subjective representation of the resulting search space (cf. Kaller, Rahm, Köstering, et al., 2011; Newell & Simon, 1972). However, despite being plausible, the attribution of internalization processes to instances of gaze before actual planning has not been shown directly but remains to be substantiated.

From a conceptual point of view, processes of internalization are closely linked to the manipulation of *Tower Configuration* as it was applied, for instance, in the fMRI study of Kaller, Rahm, Spreer, et al. (2011). In consequence, here we conducted an eye-movement study that included systematic variations of both *Search Depth* and *Tower Configuration*. The intention of the present study was to gain a better insight into the chronology of cognitive processes involved in problem solving by single-trial, microgenetic analyses (cf. Siegler & Chen, 1998) of eye movements during the planning phase. To this end, the present research addressed the question whether the same experimental manipulations that allowed us to double-dissociate activations of left and right dlPFC would also entail a corresponding dissociation in the eye-movement patterns.

Following this approach, we report converging evidence from two separate lines of research. First, the present analyses of eye-movement patterns considerably extend previous results of Kaller et al. (2009) by providing evidence for a dissociation between the temporal characteristics of cognitive processes related to experimental manipulations of *Tower Configuration* and *Search Depth*. Second, as these results entail directly testable implications for the spatio-temporal characteristics of dlPFC activation patterns during problem solving, we additionally conducted re-analyses on the fMRI data from Kaller, Rahm, Spreer, et al. (2011) that are covered in a companion paper (Ruh, Rahm, Unterrainer, Weiller, & Kaller, 2012). Taken together, the findings highlight that theoretically and behaviorally distinguishable cognitive processes during problem solving have distinct temporal characteristics and are differentially supported by left and right dlPFC.

2. Material and methods

2.1. Participants

Sixteen undergraduate students (11 women, 5 men) participated in the present study. Subjects were between 19.1 and 26.9 years of age ($M = 22.62$, SD = 2.47), right-handed and had normal or corrected-to-normal vision. None of them was under medical treatment during the examination or reported a history of psychiatric or neurological disease. Written informed consent was obtained from all of the subjects prior to the experiment. After testing, subjects received a compensation of €10 for participation. Replication data of three additional experimental groups comprising another 48 subjects in total are reported in the Supplementary materials.

2.2. Experimental procedure

Subjects were tested individually within a single session. Subjects were introduced to the Tower of London (TOL) planning task and eye-movement recordings were prepared. Prior to the actual TOL experiment, subjects practiced the handling of a three-button mouse (see below) in a set of 48 simple one- and two-ball problems (cf. also Kaller et al., 2009).

2.3. Experimental task

The Tower of London (TOL) task is one of the most commonly used disc-transfer paradigms for assessing planning ability. Originally, it was developed by Shallice (1982) to measure planning impairments in patients with frontal lesions. The TOL consists of a start state and a given goal state. Planning requirements comprise finding an efficient transformation of the start state into the desired goal state within the minimum number of moves by mentally looking ahead. The classical version of the TOL consists of three differently colored balls and three vertical rods of different heights to place the balls on. In the present study, however, participants were administered a computerized three-ball variant of the TOL with equally sized rods (cf. Ward & Allport, 1997). Due to the absence of height differences in rod sizes, the applied variant is more suitable for repeated presentation of structurally identical problems without subjects becoming aware of it (see also Kaller et al., 2009; Kaller, Rahm, Spreer, et al., 2011; Unterrainer, Rahm, Halsband, & Kaller, 2005). The start state was displayed underneath the goal state as this is the most common way of administering computerized versions of the TOL planning task. Accordingly, the data reported in the companion paper (Ruh et al., 2012) were also acquired using a vertical arrangement, with the start state being presented underneath the goal state (cf. Kaller, Rahm, Spreer, et al., 2011).

Participants were instructed to transform the start state into the goal state within the minimum number of moves. Two rules had to be taken into account: (1) only one ball could be moved at a time, and (2) a ball could not be moved if another ball was already lying on top of it. The computer program did not allow rule-incongruent moves. The balls of the start state could be moved by a sequence of two button presses using a three-button computer mouse (cf. Kaller et al., 2009; Kaller, Rahm, Spreer, et al., 2011): A ball was initially "picked up" with a press on the mouse button that corresponded one-to-one with the location of the rod on which the ball was actually placed (i.e., 'left' if the ball was placed on the left rod, 'middle' for the middle rod, and 'right' for the right rod). Accordingly, to place the ball on a certain rod, the appropriate mouse button representing the target rod had to be pressed.

Subjects were instructed to complete the task as quickly and accurately as possible. Furthermore, written and verbal instructions placed strict emphasis on planning the solution in advance of actually moving the balls. During the experiment, instructions were repeatedly presented on screen between individual blocks (see below).

The computerized version of the TOL task was implemented in the Presentation® 12.2 experimental control software package (Neurobehavioral Systems, Inc., Albany, CA). Following measures were recorded: Accuracy, initial thinking times, and movement execution times. Because the present study aimed to elucidate the temporal characteristics of planning processes, analyses were focused on initial thinking times that refer to the time window from presentation onset of a problem to the first selection of a ball.

2.4. Problem set and experimental design

The present experiment was designed to address not only the issues treated in this manuscript but also several other questions raised by the results of Kaller et al. (2009). For this reason, the experimental design comprised a full replication of the problem set that was applied in this previous study. Thus, the problem set constituted a factorial design with an equal number of one-, two-, and three-move problems with and without an intermediate move that were presented in three different combinations of start and goal configurations (cf. Kaller et al., 2009). However, as the main focus of the present manuscript was on two different kinds of processes that can be independently investigated in three-move problems only (cf. Kaller et al., 2004; Kaller, Rahm, Spreer, et al., 2011), the one- and two-move problems will be disregarded in the following.

As illustrated in Fig. B1 in Box 1, the experimental design for the analyses reported below consisted hence of a 2 × 3 factorial manipulation of two structural task parameters: *Search Depth* (two levels: with vs. without an intermediate move) and *Tower Configuration* (three levels: full start tower-partial goal tower; partial start tower – full goal tower; partial start tower – partial goal tower).

Search Depth involves different demands on mentally constructing a sequence of interdependent steps depending on whether problems do or do not require an intermediate move (cf. Kaller, Rahm, Spreer, et al., 2011). Intermediate moves are essential for an optimal solution but do not place a ball into its final goal position (Kaller et al., 2004; Spitz, Webster, & Borys, 1982). In problems requiring an intermediate move, the optimal solution can only be reliably achieved by taking into account the interdependence between the individual steps. In contrast, problems without an intermediate move may also be accomplished by pure forward processing based on simple perceptual matching-to-sample strategies (Kaller, Rahm, Spreer, Mader, & Unterrainer, 2008; Owen, 2005), that is, placing each ball in its goal position in the straightforward ordering imposed by the goal state.

By systematically varying the *Tower Configuration* of the start and goal state, the implementation of the second factor was related closely to the concept of goal hierarchy (cf. Kaller, Rahm, Spreer, et al., 2011). Goal hierarchy concerns the obviousness of goal priority as evident from the structure of the goal state (Kaller et al., 2004; Klahr & Robinson, 1981; Ward & Allport, 1997). Problems with "tower" goal states, in which all three balls are stacked on a single rod, provide an unambiguous goal hierarchy as the ball at the bottom has to be in its goal position before the ball second from the bottom, and so on. Thereby, the complete sequence in which the balls have to be placed in their goal positions is obvious from the goal state. In contrast, when the balls are distributed on two rods, as it is the case for partially ambiguous goal states, information on the ordering of the final goal moves is only partially provided. However, in three-move problems with an intermediate move, partially ambiguous goal hierarchies may be completely disambiguated by a simple consideration of the start state (cf. Kaller et al., 2004). For instance, the black ball in all three-move problems with an intermediate move in Box 1 (left side) is irrelevant to the problem's solution since it is already placed into its final position at the outset. Furthermore, in all of these problems, the gray ball is lying on top of the white ball initially. That is, by disregarding the irrelevant black dummy ball, all problems have completely unambiguous goal hierarchies. Latency effects of goal hierarchy on preplanning times in simple three-move problems (e.g., Kaller et al., 2004; Kaller, Rahm, Spreer, et al., 2011) are hence unlikely to be associated with the actual generation of the final goal moves' correct sequence but with its mere identifiability. Thus, *Configurations* of partial goal towers compared to full goal towers constitute higher demands on building up a mental representation of the encountered problem situation by identifying, matching and integrating relevant information elements between start and goal states (cf. Kaller, Rahm, Spreer, et al., 2011). Note, however, that this is valid for three-move problems only. In higher-order problems, partial *Tower Configurations* always entail partially ambiguous goal hierarchies that require searching for and generating the optimal sequence of moves (cf. Berg, Byrd, McNamara, & Case, 2010; McKinlay et al., 2008; Newman & Pittman, 2007; Ward & Allport, 1997; for discussion, see also Kaller, Rahm, Köstering, et al., 2011).

In the present study, three combinations of different full and partial *Tower Configurations* of the start and goal state were implemented (see Box 1). Together with the experimental manipulation of *Search Depth*, the resulting 2 × 3 factorial design yielded a matrix of six distinct classes of problem types. Each problem type was realized by a single structurally unique problem (Berg & Byrd, 2002) that could be solved within a minimum number of three moves and contained only one optimal path to solution.

In order to create tasks with an identical problem structure but different looks, rod arrangements and ball colors were interchanged. That is, for each of the six structurally unique problems, 12 different looking but structurally identical iso-problems were selected pseudo-randomly out of the 36 possible permutations of balls colors and rod arrangements (see Unterrainer et al., 2005). The experiment thus comprised a total of 72 trials (plus 72 one- and two-move problems). The presentation order of the problems was also pseudo-random and carefully balanced so that structurally identical problems would not occur consecutively (for further details see Kaller et al., 2009).

2.5. Recording and analysis of eye movements

During completion of the TOL planning task, eye movements were captured using an iViewX™ HiSpeed eye-tracking device (SMI SensoMotoric Instruments GmbH, Teltow/Berlin, Germany). Eye tracking was based on a dark-pupil approach during which the subject's face and eyes are illuminated with infrared (IR) light at an oblique angle and recorded concurrently with an IR-sensitive camera. Whereas eyes and face reflect the illumination, the pupil absorbs most of the IR light and appears as a high-contrast dark ellipse. Gaze locations are calculated based on prior calibrations of pupil positions in relation to defined positions on screen and further compensated for potential head movements using corneal reflections in terms of the first Purkinje image. Eye tracking was operated in monocular mode on the left eye and at a sampling rate of 350 Hz.

Subjects were seated at an eye-to-screen distance of approximately 57 cm in front of a 19-in. CRT monitor. The distance between the upper goal state and the lower start state (i.e., between the two tower bases) comprised a visual angle of approximately 16.5°. With reference to the screen center, the offset of the bottom edge of the upper and lower tower amounted +4.4° and −12.1°, respectively. Each tower's width and height spanned an angle of approximately 12° and 7.3°. The shortest distance from the screen center towards the upper and lower tower constituted hence +4.4° and −4.8°, respectively. Diameter of balls accounted for approximately 2.3°; the distance between neighboring pegs was approximately 4.5°. An overview of the layout is provided in Supplementary Fig. S1.

The iViewX™ HiSpeed eye-tracking device has an integrated chin and forehead rest with two supplementary poles that was used to stabilize subjects' heads. To keep head movements and associated needs for camera adjustments at a minimum, subjects' heads were further stabilized using an elastic tape connected to the poles.

Eye-movement recordings were synchronized with the implementation of the TOL task in the Presentation® 12.2 software package (Neurobehavioral Systems, Inc., Albany, CA) using a custom eye-tracking extension provided by SMI (SensoMotoric Instruments GmbH, Teltow/Berlin, Germany). During calibration, subjects consecutively fixated 13 stimulus points uniformly spread across the screen. For baseline recordings and continuous control of eye calibration, subjects were instructed to fixate a cross at the center of the screen for at least 500 ms before each trial. If the locus of fixation would not be detected in the expected area (visual angle of approximately 2.7°), a recalibration was initiated before the testing would continue. In addition, the experimenter monitored subjects' eye movements online on a second screen in order to initiate a recalibration whenever necessary.

Recorded eye-movement data were preprocessed using BeGaze 2.4 (SensoMotoric Instruments GmbH, Teltow/Berlin, Germany). Automatic detection of saccades was based on a peak velocity threshold of >75° per second and classification of individual fixations had to pass a minimum duration of 50 ms. Subsequent analyses of eye-movement data were processed using custom software written in MATLAB (The Mathworks Inc., Natick, MA). Based on the analyses of Kaller et al. (2009), the screen was subdivided into three Areas Of Interest (AOIs) in order to assess gaze alterations between states. Fixations on the screen with vertical gaze positions between ±2.7° (from screen center) were assigned accordingly to the central AOI whereas fixations below and above were assigned to the AOIs of the start and goal state, respectively. Saccades and resulting fixations within AOIs were aggregated and accounted for as one inspection/fixation of the corresponding state, just until the gaze position would enter another AOI. Therefore, the fixation or inspection of one state was defined as the time interval between gaze alterations crossing the border of different AOIs, whereas the term gaze shifts/gaze alteration refers merely to saccades between different states or AOIs respectively.

3. Results

For the most part, analysis procedures are analogous to Kaller et al. (2009). All reported statistics concerning analyses of variance were Greenhouse–Geisser corrected for nonsphericity where necessary (Geisser & Greenhouse, 1958).

3.1. Solution accuracy

Accuracy of the solution indicated whether the problem was correctly solved within the minimal number of moves. As expected, performance in the TOL was very accurate and close to ceiling (cf. Kaller et al., 2004, 2009; Kaller, Rahm, Spreer, et al., 2011). Across experimental manipulations of *Search Depth* and *Tower Configuration*, 88.2–90.5% and 86.5–91.7% of trials, respectively, were solved correctly in the minimum number of moves. Due to this lack of variance, accuracy was not subjected to further statistical evaluations. All subsequent analyses exclusively refer to correctly solved trials, except for the analysis of initial gaze directions.

3.2. Initial thinking time

The initial thinking time refers to the time interval between the onset of the problem presentation and the first ball movement initiated by the subject. Repeated measurements analysis of variance (RM-ANOVA) on initial thinking time revealed significant main effects of both *Tower Configuration* ($F_{(2,30)} = 20.147$, $p < .001$) and *Search Depth* ($F_{(1,15)} = 11.984$, $p = .003$) but no interaction effect ($F_{(2,30)} = 1.044$, $p = .357$). Subsequent simple-contrast analyses further revealed that initial thinking times for problems with a full start tower and a partial goal tower were significantly longer than for problems with a partial start tower and a full goal tower ($F_{(1,15)} = 12.205$, $p = .003$) whereas they were significantly shorter than for problems with a partial *Tower Configuration* in either states ($F_{(1,15)} = 4.845$, $p = .044$). As illustrated in Fig. 1, initial thinking time was increased also for problems with a high vs. a low *Search Depth* (i.e., with vs. without an intermediate move).

3.3. Gaze alternations during initial thinking time

Gaze alterations (or gaze shifts) refer to saccades between start and goal states (AOIs), whereas successive saccades within one state (AOI) were aggregated and accounted for as one fixation/inspection of the respective state. In line with previous findings of Kaller et al. (2009), problem solving during the initial thinking time was accompanied by prevailing patterns of gaze alternations which were determined substantially by the location of the first fixation. In addition, the location of the initial fixation clearly followed personal preferences as the majority of subjects (81.25%) directed their initial gaze towards the top (i.e., the goal state) in most instances (see Fig. 2).

However, the number of subsequent gaze shifts between the two tower states (AOIs) depended strongly on whether subjects had first looked at the goal state or at the start state (cf. Kaller et al., 2009). That is, if the initial fixation was directed towards the goal state (i.e., towards the top), predominant gaze patterns of four or six alternations between start and goal state were observed (inclusive of the initial gaze shift from center to goal state; Fig. 2B). In contrast, if subjects had first glanced at the start state (i.e., towards the bottom), predominant gaze patterns of three or five alternations occurred (Fig. 2C). As a consequence, final inspections before movement execution were focused almost exclusively on the start state (83.3 %) irrespective of the preceding sequence of gaze shifts between the states (cf. Kaller et al., 2009).

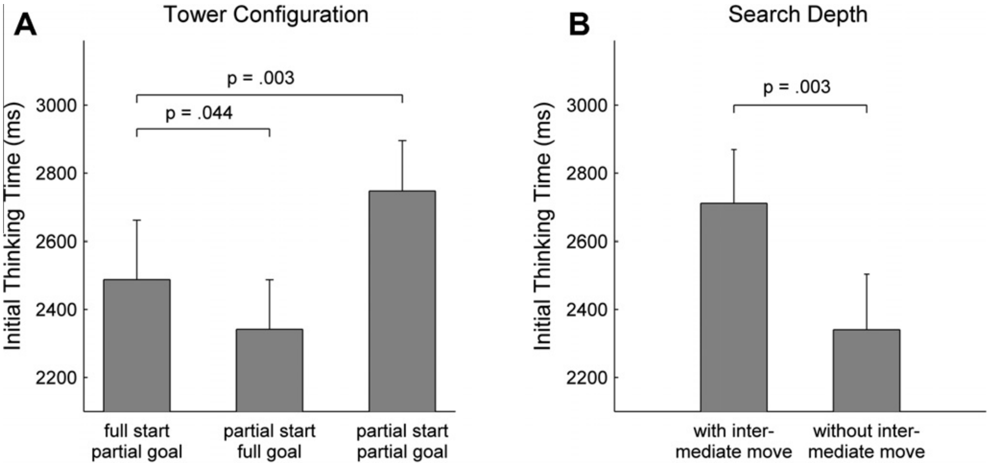

Fig. 1. Mean initial thinking times with respect to the experimental manipulations of (A) *Tower Configuration* and (B) *Search Depth*.

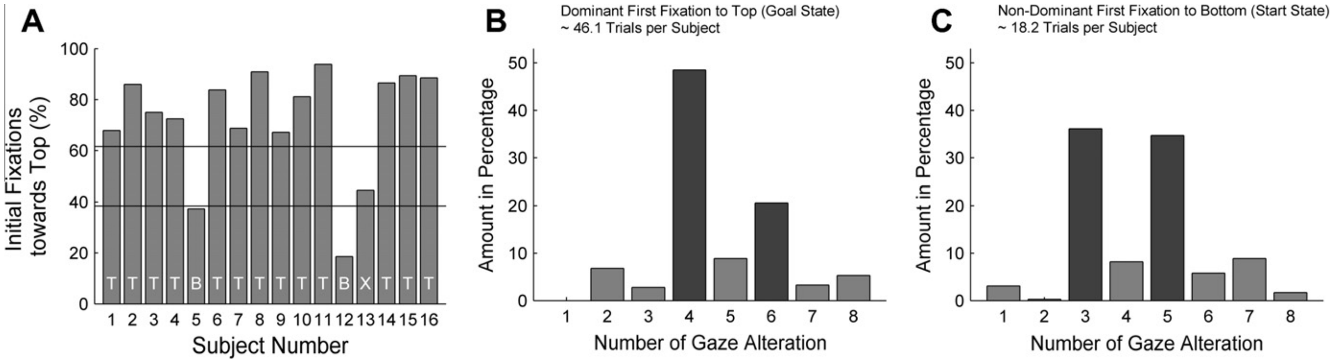

Fig. 2. (A) Proportional distribution of initial fixations for individual subjects. Classification boundaries for predominant directions are indicated by solid lines (T, top-dominant subjects; B, bottom-dominant subjects, X, subjects with evenly distributed initial fixations). Depending on whether first fixation was directed to (B) the top (goal state) or to (C) the bottom (start state), proportional distributions of gaze alternations during the initial thinking time result in a predominant gaze pattern of 4 and 6 gaze shifts (B) or 3 and 5 gaze shifts (C) in total (bars in dark gray), respectively. Note that the difference between the total number of 72 trials and the sum of average trials per subject reported in panels (B) and (C) results from the exclusion of trials that were not correctly solved within the minimum number of moves.

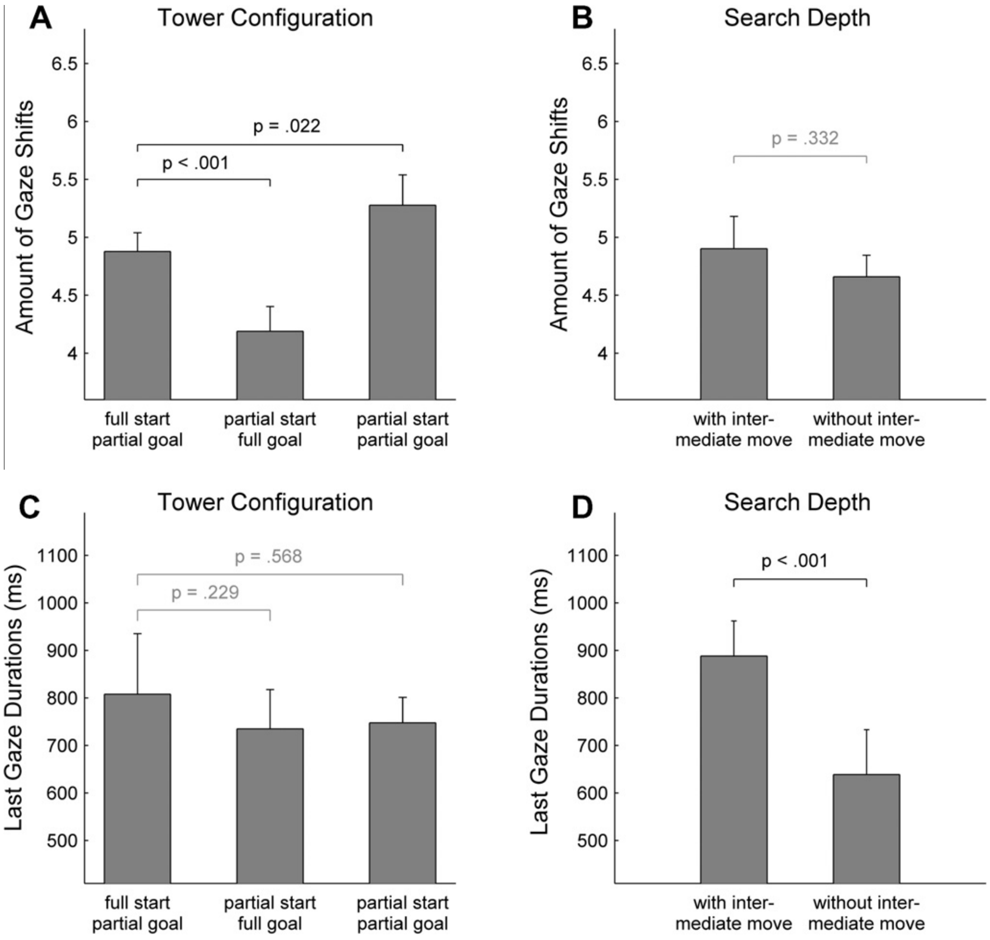

Fig. 3. Dissociation between the two experimentally manipulated task parameters and two different eye-movement measures. More specifically, panels illustrate the number of gaze shifts (A and B) and the duration of the last fixation during the initial thinking time (C and D). In addition, left (A and C) and right panel positions (B and D) refer to experimental manipulations on *Tower Configuration* (full start–partial goal/partial start–full goal/partial start–partial goal) and *Search Depth* (with/without an intermediate move), respectively.

In order to assess the influence of different cognitive aspects of problem solving during the initial thinking time, the number of gaze shifts was entered as a dependent variable in a two-factorial RM-ANOVA with the independent variables *Search Depth* and *Tower Configuration*. As the majority of subjects directed their initial gaze mainly towards the top (see above), the analysis was restricted to the trials starting with a gaze fixation of the upper goal state. Results revealed a significant main effect for *Tower Configuration* ($F_{(2,28)} = 35.280$, $p < .001$) but neither a main effect for *Search Depth* ($F_{(1,14)} = 1.009$, $p = .332$) nor an interaction effect

Table 1
Influence of Tower Configuration and Search Depth on fixation durations of alternating inspections of start and goal state during the initial thinking time.

Inspection	Search depth				Tower configuration					Interaction	
	Descriptive statistics *M*(SEM)		Inference statistics		Descriptive statistics *M*(SEM)			Inference statistics		Inference statistics	
	With i-move	Without i-move	*F*	*p*	Full partial	Partial full	Partial partial	*F*	*p*	*F*	*p*
1st (goal)	402.50 (60.16)	404.15 (65.97)	0.014	.907	395.91 (55.93)	421.01 (71.35)	393.06 (66.08)	0.613	.514	4.780	.025
2nd (start)	379.94 (25.96)	371.02 (33.53)	0.209	.658	367.14 (25.41)	353.45 (25.70)	405.84 (46.01)	1.462	.257	0.858	.398
3rd (goal)	472.56 (35.17)	478.28 (35.09)	0.044	.839	477.12 (45.74)	493.50 (24.69)	455.63 (35.97)	0.844	.425	1.346	.282
4th (start)	888.20 (73.72)	638.82 (94.56)	39.601	<.001	807.87 (127.50)	735.10 (82.61)	747.58 (53.55)	0.545	.497	0.246	.738

N.B. Fixation durations are reported in milliseconds. Abbreviation: i-move, intermediate move.

($F_{(2,28)} = .059$, $p = .923$). The underlying patterns of gaze shifts are displayed in Fig. 3A and B. Subsequent simple contrast analyses revealed that problems with a full start tower and a partial goal tower were associated with more gaze shifts than problems with a partial start tower and a full goal tower ($F_{(1,14)} = 108.380$, $p < .001$) whereas it took fewer gaze shifts than for problems with a partial *Tower Configuration* in either states ($F_{(1,14)} = 6.582$, $p = .022$).

3.4. Fixation duration during initial thinking time

Fixation duration or inspection duration refers to the time interval between gaze alterations, i.e., the dwell time on the AOIs corresponding to start and goal state. Besides the number of gaze shifts during the initial thinking time, the duration of the respective inspections of the start state and the goal state may also yield insights into the cognitive processes associated with planning (cf. Kaller et al., 2009). Because independence of subsequent gaze shifts might not be assumed, multiple univariate analyses would be susceptible to distortions by mutual dependencies within gaze patterns. The gaze durations of individual inspections were therefore entered as multiple dependent variables into a multivariate ANOVA with repeated measurements (RM-MANOVA) with factors *Search Depth* and *Tower Configuration*. Hence, duration data had to be separately analyzed for different patterns of gaze shifts. In order to ensure a sufficient number of observations, analyses were applied to the most frequent pattern across subjects, namely four gaze shifts after an initial fixation toward the goal state at the top (cf. Fig. 2B; see also Kaller et al., 2009). As evident from Table 1 as well as from Fig. 3C and D, results revealed a main effect of *Search Depth* only for the duration of the last inspection of the start state before movement execution ($F_{(1,10)} = 39.601$, $p < .001$), but not for any of the preceding inspections (all $F_{(1,10)} < .210$, all $p > .657$). In contrast, no main effect of *Tower Configuration* was observed neither for the final inspection of the start state ($F_{(2,20)} = .545$, $p = .497$) nor for any of the preceding inspections (all $F < 1.463$, all $p > .256$). A significant interaction effect between *Search Depth* and *Tower Configuration* was found for the duration of the first inspection ($F_{(2,20)} = 4.780$, $p = .025$) but not for any of the other inspections (all $F < 1.347$, all $p > .281$).

4. Discussion

In the present study, experimental manipulations of two different aspects of problem solving disclosed a process-contingent dissociation of cognitive processing during the initial thinking time: Task demands imposed by different variations of *Tower Configuration* were reflected solely by changes in the total number of gaze alternations between start and goal state. In contrast, task demands imposed by variation of *Search Depth* exclusively influenced the duration of the last fixation before movement execution. Hence, present results complement previous fMRI findings of a double dissociation between left and right dlPFC function (Kaller, Rahm, Spreer, et al., 2011) with novel evidence for a corresponding dissociation in eye-movement parameters during planning as a marker for the associated cognitive processes. Most importantly, present data provide the first empirical support for a temporal and functional separation of internalization and actual planning as dissociable components in the course of problem solving – a separation that was already argued for based on plausibility in an earlier study (Kaller et al., 2009). Present data also fully replicate previous results of Kaller et al. (2009) demonstrating that variations in *Search Depth* (see also Footnote 2) affected particularly the duration of the last fixation before movement execution. Beyond that, the present findings of dissociating eye-movement patterns following systematic variations of *Search Depth* and *Tower Configuration* were also replicated in additional experimental conditions (see Supplementary materials).

In detail, the present results constitute an important extension of the recently reported double dissociation between the function of left and right dorsolateral prefrontal cortex (dlPFC) and *Tower Configuration* and *Search Depth*, respectively (Kaller, Rahm, Spreer, et al., 2011). In particular, the observed dissociation between different temporal characteristics of eye-movement patterns and the two problem parameters (cf. Fig. 3) complement previous functional interpretations of dissociating dlPFC activation: Based on the experimental operationalization of *Tower Configuration*, left dlPFC function was associated with "building up a mental representation of the encountered problem situation in terms of identifying, matching, and integrating relevant information pieces between start and goal state" (p. 314, Kaller, Rahm, Spreer, et al., 2011; cf. also Grafman, Spector, & Rattermann, 2005). The present understanding of internalization processes not only refers to establishing a visual representation but also assumes to include the generation of a subjective representation of the search space that is not necessarily an exhaustive mapping of the objective problem space (cf. Newell & Simon, 1972). In line with this, higher demands imposed by manipulations of *Tower Configuration* led to an increasing number of gaze shifts between start and goal state but did not affect the duration of the last inspection of the start state before movement execution (Fig. 3). In contrast, right dlPFC activation following higher demands on *Search Depth* was related to processes of "integrating interdependent information into a coherent action sequence", that is, actual planning (p. 314–315, Kaller, Rahm, Spreer, et al., 2011; cf. also Grafman et al., 2005). Again in support of the previous functional interpretations, here manipulations of *Search Depth* were shown to affect the duration of the last inspection of the start state before movement execution (see also Kaller et al., 2009) but not the number of gaze shifts between start state and goal state.

The present results are in accord with common theoretical frameworks of cognitive processing during problem solving that assume a serial order of internalization and planning (Hodgson et al., 2000; Newell & Simon, 1972; Ward & Morris, 2005). Yet, although a cognitive function might reliably coincide with a specific eye-movement pattern, it is impossible to determine exactly how the underlying cognitive functions operate merely on the basis of observed gaze data (cf. Droll & Hayhoe, 2007; Hayhoe, 2004). Therefore, the question remains of actually how information is encoded, stored and modulated in visuo-spatial problem solving tasks. A crucial point in this regard is the issue of limitations in working memory capacity. As planning ahead is inevitable for efficient problem solving in the TOL task, a just-in-time method of gathering information through strategic fixations is not sufficient. Instead, subjects have to rely on working memory in order to generate and manipulate an internal model of the encountered problem. However, given that planning seems to occur during the last inspection of the start state immediately before movement execution, it may be beneficial and less resource-consumptive if subjects do not internally represent the start state but only the goal state. Consequently, internal representations may be restricted to the goal state or even fragments of it, while the fixation of gaze on the start state during planning may allow to reduce working memory load and instead to acquire information about the start state just in time. Several studies on eye movements indicate that subjects tend to keep memory load substantially below posited constraints (Ballard, Hayhoe, Pook, & Rao, 1997; Droll & Hayhoe, 2007; Droll, Hayhoe, Triesch, & Sullivan, 2005). That is, in order to reduce working memory load, information is not stored in internal representations comprehensively, but gathered just in time through strategic shifts of gaze. In that sense, information is not internalized but remains external and is used as a form of external memory (Ballard, Hayhoe, & Pelz, 1995; Ballard et al., 1997; Hayhoe, Bensinger, & Ballard, 1998; Hoffman, Landau, & Pagani, 2003; O'Regan, 1992; O'Regan & Noe, 2001).

With respect to the present study, the use of an external representation of the start state is suggested by the significant increases in both initial thinking times (Fig. 1A) as well as the number of gaze alternations (Fig. 3A) in problems with full start – partial goal towers compared to problems with partial start – full goal towers. From their purely physical layout, these two types of problems feature an identical amount and quality of information to be internalized. That is, each of these two problem types consists of both a partial and a full tower – only the assignments to the start and goal states are inverted. Thus, assuming that building up a mental representation of start and goal configurations cannot account for the observed differences, present data may suggest that subjects keep information about start states (or at least some parts of it) externalized while using the internalized goal state as reference for mental generation and evaluation of alternative move sequences.

Given the cognitive demands imposed by a visuo-spatial problem solving task such as the TOL, the establishment of an internal model as a mere visual mapping appears unlikely. Speculating about the nature of an internal representation, an internal model most likely includes abstracted information not only about individual ball positions but also about spatial relations and imposed move restrictions. In other words, planning requires an understanding of the facts that constitute the very problem (i.e., the given options and restrictions). Consequently, an internal representation most likely comprises a subjective mapping of the search space (cf. Newell & Simon, 1972). This is also suggested by comparing partial goal tower problems with different start states: In problems with a partial tower in the goal state, subjects exhibited more gaze alternations between states and showed longer initial thinking times when the start state was another partial tower as compared to when it was a full start tower (Figs. 1A and 3A). That is, the redundancy of gaze alternations between states may indicate that the generation of internal representations also includes comparing and matching of individual balls and their positions between start and goal state so as to identify the particular differences between states and to define the balls relevant for goal succession. If the focus of internalizing was placed on the goal state (see above), the observed differences in gaze alternations and initial thinking times cannot be explained by simply establishing a mental copy of the external goal state because both types of problems (i.e., with partial start – partial goal and full start – partial goal towers) feature identical configurations of partial towers in the goal state. These specific differences in gaze alternation patterns between problems with a partial goal tower and either partial or full start towers can, however, be explained by considering the problems' features. To gather information about relations amongst balls, it is likely that redundant gaze alternations between states also entail processes of comparing, identifying and encoding the balls' entities along with their allocations and positions on individual rods. In this respect, matching information between two partial towers is likely to be more demanding and resource-consumptive than between a partial goal and a full start tower. For instance, spatial relations of balls on a full tower can be encoded in one dimension (i.e., the relative positions of balls on a single rod) whereas two dimensions are required in a partial tower condition (i.e., the balls' allocations to individual rods, and the balls' respective positions on these rods). Thus, if the fixation alternations reflect internalization processes, and if it is indeed the goal state that is internalized, these findings suggest that its internal representation depends on comparisons with features of the start state.

Although previous research conceived internalization processes mostly as visual mappings of the external world instead of an abstract representation, some findings are also relevant here. For instance, studies on simple visuo-motor tasks requiring subjects to copy meaningless models consisting of colored blocks revealed that repeated gaze alternations between model and work space served the separate acquisition of different target features such as color and location (Ballard et al., 1995; see also Ballard et al., 1997). In addition, an eye-movement study on problem solving in geometry found that visual images are added to memory by redundant oculo-motor scanning (Epelboim & Suppes, 2001). Taking into account that in the TOL the last fixation during initial thinking time is assigned to actual planning (Kaller et al., 2009) and that at least some internal model of the goal state has to be established in advance, we conclude that the sequence of repeated gaze alternations between start and goal state serves to identify and encode pieces of information relevant and necessary for subsequent planning.

Another important aspect of the present study concerns the replication of the findings on actual planning and manipulations of *Search Depth* from a previous eye-movement study (see Kaller et al., 2009). This is particularly intriguing with respect to the vast differences in methodical approaches and physical layouts between studies. In the previous study Kaller et al. (2009), eye movements were captured by electrooculography (EOG) that – compared to the use of infrared-based eye-tracking in the present study – has a considerably lower spatial resolution. Even more importantly, the substantially more widely separated spatial layout in the previous study with its increased distances between start and goal state incorporated relatively high costs for programming and executing saccades between states and hence may have purposely precluded parafoveal processing of information. In addition, problem states were presented horizontally at the left and right margins of the screen instead of one upon the other and at a close distance. Moreover, heights and widths of start and goal states in the previous study were less than half the sizes used in the present study. That is, the results of Kaller et al. (2009) might have been induced by the experimental design rather than constituting a natural feature of problem solving. Nevertheless, fully replicating the result that variations in *Search Depth* exclusively determined the duration of the last fixation before movement execution despite the methodical differences between studies renders it a reliable and stable phenomenon.

Finally, the present study on dissociable phases of cognitive processing during problem solving followed a parsimonious approach by focusing on simple three-move problems only. From an experimental point of view, this approach bears considerable advantages as it allows identification of basic principles of higher-order cognition based on a sufficient number of observations while limiting fluctuations in information processing due to inter- and intraindividual variations in strategic accounts, capacity constraints, and other sources of task impurity (Frith, Gallagher, & Maguire, 2004). Yet, it has to be taken into account that the generalizability of the present approach may have its limitations. Future studies will have to address the extent to which regularities found for three-move problems can be transferred to planning and problem solving in more demanding higher-order problems. For instance, Berg et al. (2010) have reported that start state and goal state configurations were important determinants of problem difficulty, with both states significantly contributing to different response measures. In this respect, a major issue will concern disentangling the different cognitive demands that are associated with the internalization from those that are related to the ambiguity with which the sequence of the final goal moves can be derived from the goal state (“goal hierarchy”; see also Section 2 and Kaller et al., 2004; Kaller, Rahm, Köstering, et al., 2011). Future research will also have to address the question of how adequately cognitive processing in well-structured disc-transfer tasks like the TOL reflects processing in real-world planning situations (cf. Burgess, Simons, Coates, & Channon, 2005; Goel, Grafman, Tajik, Gana, & Danto, 1997; Goel, Pullara, & Grafman, 2001). Furthermore, the present approach may also be extended to clinical populations. In previous studies, detailed analyses of eye-movement patterns in the Tower of London task have already proven fruitful for enhancing the understanding of planning-related deficits, for instance, in Parkinson’s disease and schizophrenic patients (see Hodgson, Tiesman, Owen, & Kennard, 2002; Huddy et al., 2007).

Taken together, growing evidence suggests temporally and functionally separable phases of information processing during problem solving, with internalization processes being followed by actual planning processes (cf. Hodgson et al., 2000; Kaller et al., 2009). Moreover, a second line of research shows that the prefrontal correlates underlying these two cognitive processes double-dissociate across hemispheres as processes of internalization and planning are differentially driven by left and right dlPFC, respectively (Kaller, Rahm, Spreer, et al., 2011). Thus, when combining these two findings, one might expect that processes of internalization and planning should also be – at least to some extent – distinguishable in time and space in terms of neural activation patterns. In a companion paper (Ruh et al., 2012), we report a spatio-temporal re-analysis of the original fMRI data from Kaller, Rahm, Spreer, et al. (2011) that reveals substantial support for this proposal.

Acknowledgments

This work was supported by grants of the German Federal Ministry of Education and Research (BMBF; Grant Number 01GW0710) and the Freiburg University Scientific Society (“Wissenschaftliche Gesellschaft Freiburg i.Br.”). We thank Lena Köstering, Katharina Heinze, Sandra Loosli and the authors of the companion paper (Ruh et al., 2012) for invaluable discussion of an earlier version of this manuscript.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bandc.2012.05.003.

References

Ballard, D. H., Hayhoe, M. M., & Pelz, J. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience, 7*, 66–80.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences, 20*, 723–767.

Berg, W. K., & Byrd, D. L. (2002). The Tower of London spatial problem-solving task: Enhancing clinical and research implementation. *Journal of Clinical and Experimental Neuropsychology, 24*, 586–604.

Berg, W. K., Byrd, D. L., McNamara, J. P., & Case, K. (2010). Deconstructing the tower: Parameters and predictors of problem difficulty on the Tower of London task. *Brain and Cognition, 72*, 472–482.

Burgess, P., Simons, J., Coates, L., & Channon, S. (2005). The search for specific planning processes. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 199–227). Hove: Psychology Press.

Droll, J., & Hayhoe, M. M. (2007). Trade-offs between working memory and gaze. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1352–1365.

Droll, J. A., Hayhoe, M. M., Triesch, J., & Sullivan, B. T. (2005). Task demands control acquisition and storage of visual information. *Journal of Experimental Psychology, 31*, 1416–1438.

Epelboim, J., & Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Research, 41*, 1561–1574.

Frith, C., Gallagher, H., & Maguire, E. (2004). Mechanisms of cognitive control. In R. S. J. Frackowiak, K. J. Friston, C. D. Frith, R. J. Dolan, C. J. Price, S. Zeki, J. Ashburner, & W. Penny (Eds.), *Human brain function* (pp. 329–362). London, UK: Academic Press.

Geisser, S., & Greenhouse, S. W. (1958). An extension of Box’s results on the use of the *F*-distribution in multivariate analysis. *Annals of Mathematical Statistics, 29*, 885–891.

Goel, V., & Grafman, J. (1995). Are frontal lobes implicated in planning functions: Re-examining the data from the Tower of Hanoi. *Neuropsychologia, 33*, 623–642.

Goel, V., Grafman, J., Tajik, D., Gana, S., & Danto, D. (1997). A study of the performance of patients with frontal lobe lesions in a financial planning task. *Brain, 120*, 1805–1822.

Goel, V., Pullara, D., & Grafman, J. (2001). A computational model of frontal lobe dysfunction: Working memory and the Tower of Hanoi task. *Cognitive Science, 25*, 287–313.

Grafman, J., Spector, L., & Rattermann, M. J. (2005). Planning and the brain. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 181–198). Hove: Psychology Press.

Hayhoe, M. M. (2004). Advances in relating eye movements and cognition. *Infancy, 6*, 267–274.

Hayhoe, M. M., Bensinger, D. G., & Ballard, D. H. (1998). Task constraints in visual working memory. *Vision Research, 38*, 125–137.

Hodgson, T. L., Bajwa, A., Owen, A. M., & Kennard, C. (2000). The strategic control of gaze direction in the Tower of London task. *Journal of Cognitive Neurosciences, 12*, 894–907.

Hodgson, T. L., Tiesman, A., Owen, A. M., & Kennard, C. (2002). Abnormal gaze strategies during problem solving in Parkinson’s disease. *Neuropsychologia, 40*, 411–422.

Hoffman, J. E., Landau, B., & Pagani, B. (2003). Spatial breakdown in spatial construction: Evidence from eye fixations in children with Williams syndrome. *Cognitive Psychology, 46*, 260–301.

Huddy, V. C., Hodgson, T. L., Kapasi, M., Mutsatsa, S. H., Harrison, I., Barnes, T. R., et al. (2007). Gaze-strategies during planning in first-episode psychosis. *Journal of Abnormal Psychology, 116*, 589–598.

Kaller, C. P., Rahm, B., Bolkenius, K., & Unterrainer, J. M. (2009). Eye movements and visuospatial problem solving: Identifying separable phases of complex cognition. *Psychophysiology, 46*, 818–830.

Kaller, C. P., Rahm, B., Köstering, L., & Unterrainer, J. M. (2011). Reviewing the impact of problem structure on planning: A software tool for analyzing tower tasks. *Behavioural Brain Research, 216*, 1–8.

Kaller, C. P., Rahm, B., Spreer, J., Mader, I., & Unterrainer, J. M. (2008). Thinking around the corner: The development of planning abilities. *Brain and Cognition, 67*, 360–370.

Kaller, C. P., Rahm, B., Spreer, J., Weiller, C., & Unterrainer, J. M. (2011). Dissociable contributions of left and right dorsolateral prefrontal cortex in planning. *Cerebral Cortex, 21*, 307–317.

Kaller, C. P., Unterrainer, J. M., Rahm, B., & Halsband, U. (2004). The impact of problem structure on planning: Insights from the Tower of London task. *Cognitive Brain Research, 20*, 462–472.

Klahr, D., & Robinson, M. (1981). Formal assessment of problem-solving and planning processes in preschool children. *Cognitive Psychology, 13*, 113–148.

McKinlay, A., Kaller, C. P., Grace, R. C., Dalrymple-Alford, J. C., Anderson, T. J., Fink, J., et al. (2008). Planning in Parkinson’s disease: A matter of problem structure? *Neuropsychologia, 46*, 384–389.

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

Newman, S. D., & Pittman, G. (2007). The Tower of London: A study of the effect of problem structure on planning. *Journal of Clinical and Experimental Neuropsychology, 29*, 333–342.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology, 46*, 461–488.

O'Regan, J. K., & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*, 939–1031.

Owen, A. M. (2005). Cognitive planning in humans: New insights from the Tower of London (TOL) task. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 135–152). Hove: Psychology Press.

Ruh, N., Rahm, B., Unterrainer, J. M., Weiller, C., & Kaller, C. P. (2012). Dissociable stages of problem solving (II): First evidence for process-contingent temporal order of activation in dorsolateral prefrontal cortex. *Brain and Cognition, 80*, 170–176.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society London B, 298*, 199–209.

Siegler, R. S., & Chen, Z. (1998). Developmental differences in rule learning: A microgenetic analysis. *Cognitive Psychology, 36*, 273–310.

Spitz, H. H., Webster, N. A., & Borys, S. V. (1982). Further studies of the Tower of Hanoi problem-solving performance of retarded young adults and nonretarded children. *Developmental Psychology, 18*, 922–930.

Unterrainer, J. M., Rahm, B., Halsband, U., & Kaller, C. P. (2005). What is in a name: Comparing the Tower of London with one of its variants. *Cognitive Brain Research, 23*, 418–428.

Ward, G., & Allport, A. (1997). Planning and problem-solving using five-disc tower of London task. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 50A*, 49–78.

Ward, G., & Morris, R. G. (2005). Introduction to the psychology of planning. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 1–34). Hove: Psychology Press.