Doctoral thesis submitted to

the Faculty of Behavioural and Cultural Studies

of Heidelberg University

in partial fulfillment of the requirements of the degree of

doctor philosophiae (Dr. phil.)

in Psychology

Planning Ability in Schizophrenia

by

Daniel V. Holt

2012

Dean: Prof. Dr. Klaus Fiedler

Advisors: Prof. Dr. Joachim Funke

Prof. Dr. Matthias Weisbrod

Acknowledgements

I would like to thank my advisor Prof. Joachim Funke for being supportive of the research projects presented in this thesis throughout and for always being available to listen to and discuss questions whenever it was necessary.

I would also like to thank my collaborators and co-authors who contributed to the research presented in this thesis: Katlehn Rodewald, Mirjam Rentrop, Prof. Matthias Weisbrod, Jürgen Wolf, Sven Brüssow, Matthias Backenstrass, Daniela Roesch-Ely and Stefan Kaiser.

Thank you as well to the student assistants and clinical staff who supported us during data collection.

Special thanks to Andreas Fischer and Helen Fischer for various helpful comments on earlier versions of this thesis and to Sarah for proof-reading it.

And of course, I would like to thank the patients and other volunteers who took the time to participate in these studies. I its my sincere hope and aim that the advances in this field of research will contribute towards improving clinical practice.

Finally, I would like to thank Sarah for her patience and Elias and Raphael for just being there. Now we finally have time to spend Sunday afternoons together just hanging out!

Contents

1 Introduction

"*I love it when a plan comes together.*" – John 'Hannibal' Smith, *The A-Team*

Planning is central to many human activities, from making a cup of tea, over planning a holiday, to drafting a complex business plan. It can take milliseconds for movement and speech planning or years and decades for planning a space mission or economic reforms in a state. Planning can take place in someone's mind or involve external aids, such as maps, blueprints and written plans. It can be carried out by a single person or by groups and institutions, such as the management team of a company. The adaptive value of planning ability may be so fundamental that it is not even exclusive to humans but also found in a rudimentary form in other species with sufficiently developed cognitive systems, such as scrub jays or chimpanzees (Raby, Alexis, Dickinson, & Clayton, 2007; Osvath & Osvath, 2008). In the present thesis I will adopt a cognitive perspective of planning as a mental process for formulating a course of action to reach one or more goals (Friedman & Scholnick, 1997). In particular, the focus of this thesis lies on the measurement of planning ability in psychiatric patients with schizophrenia, combining a problem solving perspective with a psychometric approach to assessing executive functions.

1.1 Planning and schizophrenia

Schizophrenia is a severe psychiatric disorder with a lifetime prevalence of about 0.7% (McGrath, Saha, Chant, & Welham, 2008). Its key symptoms include disorganized speech and thinking, visual or auditory hallucinations, paranoid or bizarre delusions and reduced emotional responsiveness (American Psychiatric Association, 2000). In addition to affecting the subjectively perceived quality of life, these symptoms can notably impair social and occupational functioning and the general ability for independent living. While symptoms such as hallucinations and delusions are among the better known signs of schizophrenia, it has meanwhile been well established that the disorder adversely affects a wide range of cognitive abilities, from learning and memory over verbal ability to executive functions (Heinrichs and Zakzanis, 1998). Executive functions in general have been shown to be predictive of everyday performance in schizophrenia (Green, 1996; Green, Kern, Braff, & Mintz, 2000; Velligan, Bow-Thomas, Mahurin, Miller, &

Halgunseth, 2000), and this may be particularly the case for planning ability, which is integral to many everyday activities (Miller, Galanter & Pribram, 1960).

Planning impairments in schizophrenia have been described in detailed analyses of activities of daily living (Semkovska, Bédard, Godbout, Limoge, & Stip, 2004; Seter, Giovannetti, Kessler, & Worth, 2011) and there is initial evidence for the association of planning ability and functional outcome (Katz, Tadmor, Felzen, & Hartman-Maeir, 2007; Wykes, Reeder, Huddy, Taylor, Wood, Ghirasim, et al., 2012). Neuropsychological studies also consistently show a replicable planning deficit in patients measured by standardized planning tasks, usually the Tower of London or its variants (e.g., Chan, Chen, Cheung, Chen, & Cheung, 2004; Marczewski, Van der Linden, & Larøi, 2001; Morris, Rushe, Woodruffe, & Murray, 1995). Even when matching comparison groups for general intellectual ability and controlling for the effects of psychomotor retardation or deficits in spatial working memory patients with schizophrenia showed a clear planning deficit (Hutton, Puri, Duncan, Robbins, Barnes, & Joyce, 1998; Pantelis, Barnes, Nelson, Tanner, Weatherley, Owen, et al., 1997; Rushe, Morris, Miotto, Feigenbaum, Woodruff, & Murray, 1999). Indeed, Morris et al. (1995) suggest that a specific deficit in planning ability may be a characteristic feature of the cognitive impairments in schizophrenia. This view is also supported by Johnstone and Frith (1996), who argue that the difficulties in generating plans, goals and intentions are related to other cognitive and motivational impairments in schizophrenia.

One mechanism that has been proposed as a cognitive explanation of this planning deficit is an impairment of action selection at the level of the supervisory attentional system and its fronto-striatal neural underpinnings (Norman & Shallice, 1986; Robbins, 1990; Frith, 1987). More specifically, the effect may be caused by a deficit in response inhibition, which has been well established in schizophrenia for other tasks (e.g., Fuller, Frith, & Jahanshahi, 2000; Nathaniel-James, Brown, & Ron, 1996). In planning a deficit in response inhibition could induce patients to act prematurely before planning is completed. This fits with the commonly observed pattern that planning times for patients are comparable to control groups, but their subsequent planning and execution times and error rates are increased, which can be interpreted as an indication for insufficient planning (e.g., Morris et al., 1995; Pantelis et al., 1997; Hutton et al., 1998; Chan et al., 2004). However, this line of evidence is not conclusive, since errors could also be caused by problems in monitoring the execution of a plan. Moreover, a detailed analysis by Marczewski et al. (2001) showed that patients were able to inhibit perceptually cued but incorrect moves just as well as healthy control participants. At present the evidence for a simple response inhibition deficit as an explanation for impaired planning ability – although in principle plausible – seems equivocal.

The explanation preferred in this thesis is based on a different proposal made by Bustini, Stratta, Daneluzzo, Pollice, Prosperini, and Rossi (1999), who explained deficits in planning and problem solving in schizophrenia by reference to the internal representation of context (Cohen and Servan-Schreiber, 1992; Cohen, Barch, Carter & Servan-Schreiber, 1999). The internal representation of context includes information about task rules, instructions, prior stimuli or results of previous processing steps that are relevant for an appropriate response. Cohen and Servan-Schreiber (1992) showed that a deficit of this type can explain various cognitive deficits in schizophrenia and how it may be related to disordered dopaminergic activity in the prefrontal cortex. This could explain why tasks with high demands in representing and updating contextual information – such as planning and problem-solving tasks – are disproportionally affected compared to tasks involving a relatively static context, e.g., sustained attention tasks.

This explanation also matches the response time and error patterns of the planning studies cited above (Morris et al., 1995; Pantelis et al., 1997; Chan et al., 2004), as difficulties in using and updating contextual information should become particularly pertinent during plan execution, resulting in more errors and slower execution, whereas the initial planning phase is less affected. Beyond impaired planning ability, an impaired representation of internal context could also account for the increased number of violations of task rules and difficulties in rule-finding tasks such as the Wisconsin Card Sorting Task, which have high requirements for the internal representation of context (Bustini et al., 1999). As this theory parsimoniously explains a range of phenomena in planning and problem solving, it is the preferred explanation in this thesis. This is particularly relevant for Manuscript 3, which contrasts the abilities of patients with schizophrenia with another group with depression and a healthy control group.

Furthermore, this account is compatible with another phenomenon in the planning literature in schizophrenia, the “complexity effect”, i.e., the observation that planning impairments in schizophrenia increase disproportionally with task complexity (e.g., Hilti, Delko, Orosz, Thomann, Ludewig, Geyer, et al., 2010; Morris et al., 1995; Marczewski et al., 2001). Arguably, more complex tasks pose higher demands in terms of internal representation of context, e.g., the number of intermediate steps required. More generally, complex tasks pose higher demands with regard to selecting, efficiently representing, organizing, and updating task-related information. Rather than being a simple deficit in the ability to perform cognitive “look ahead” operations, a planning deficit in schizophrenia may therefore be related to an overarching deficit in information selection, strategy formation and action monitoring (cf. Burgess, Simons, Coates, & Channon, 2005). The relatively open nature of planning tasks, i.e., the need to structure a situation and develop a strategy, may make them more sensitive to deficits of this type than other tasks. This perspective will be taken up in chapter 4, which will present a supplementary cognitive process analysis accompanying the psychometric approaches to measuring planning ability.

In sum, there is a robust empirical evidence that planning ability – as measured by classical neuropsychological planning tests – is impaired in schizophrenia and initial findings support the functional relevance of such a planning deficit. Furthermore, there are indications that a planning deficit may be particularly characteristic for schizophrenia and that it may be related to aspects of the underlying psychopathology. Before returning to the question of how planning ability can be measured in patients, I will briefly outline the cognitive and neuropsychological foundations of planning ability.

1.2 A cognitive perspective on planning

Miller, Galanter and Pribram (1960) firmly put planning on the agenda of cognitive psychology as a central aspect of human cognitive activity that serves to organize thought and behavior. According to their definition, a plan is (Miller et al., 1960, p. 16):

> *… any hierarchical process in the organism that can control the order in which a sequence of operations is to be performed.*

Miller et al. (1960) take a computational perspective, describing plans as program-like structures that are created and executed by a “human processor”. They conceive of plans as hierarchical, with superordinate and subordinate plans at different levels of abstraction, e.g., a subplan for looking up train times as part of a superordinate plan to travel to another city. The execution of plans is guided by nested conditional feedback loops named *test-operate-test-exit units*. These conditional loops control the flow of execution of a plan by testing whether a particular subgoals has been reached and passing on control to other parts of the plan accordingly. In principle this mechanism is powerful enough to model any form of behavior control that can be computationally described.

About a decade later this computational perspective was greatly expanded in one of the most detailed computational analyses of human cognition: Newell & Simon’s (1972) classic *Human Problem Solving* located planning within a comprehensive cognitive problem solving framework. Indeed, the terms planning and problem solving are sometimes used almost synonymously in the literature (cf. Funke & Glodowski, 1990). Although the two processes are certainly related and can be nested within each other (e.g., discovering an unforeseen problem while making a plan) they are not strictly identical. According to a widely used definition, problem solving involves the transformation of an initial state of a situation to a particular goal state by overcoming a barrier that requires the application of non-routine behavior and strategies (cf. Frensch & Funke, 1995). Planning, however, neither necessarily involves a barrier (e.g., a simple plan for making a cup of tea) nor does it require overt behavior to transform a situation. Conversely, some problems can be solved without planning, e.g., by trial and error. However, in practice planning and problem solving are often intertwined, and the present thesis focuses on planning as part of a problem solving process, i.e., non-routine planning. How closely planning and problem-solving are related is also evident in the practice of planning research and assessment, where the Tower of Hanoi and its variants – a classical problem solving task – is one of the most widely used paradigms (Ward & Morris, 2005).

While planning had implicitly been part of problem solving theories before, Newell and Simon (1972) "removed the magic" from human planning by explicating the computational procedure in such detail that human planning could, in principle, be simulated on a computer. This work also laid the foundation for some of the earliest computer-implemented cognitive models of planning and problem solving, e.g., for the Tower of Hanoi planning problem, solving logic problems, or playing chess (e.g., Anzai & Simon, 1979; Newell & Simon, 1972). This line of research has been extended by Anderson and colleagues, who constructed comprehensive computational models of human planning in a cognitive architecture that predicts subtle aspects of planning performance and characteristic errors in planning (e.g., Anderson & Douglass, 2001). The validity of this approach has been confirmed by combining these models with modern neuroimaging techniques (e.g., Anderson, Albert, & Fincham, 2005). Not only was the model able to accurately predict behavioral data, but it also predicted non-linear activations pattern in particular brain regions with surprising precision. The cognitive perspective adopted in this thesis follows the tradition of Miller et al. (1960), Newell & Simon (1972), and Anderson (2001, 2005), in viewing planning as a computational process within a problem solving framework.

1.3 An executive functions perspective on planning

While the cognitive psychology perspective outlined above strongly emphasizes the cognitive processes involved in planning, the executive functions perspective also considers the role of planning as measurable ability. This perspective is captured by the definition of Caroll (1988, p. 848), who emphasizes the ability aspect of planning[1]:

> *Planning (PL) is the ability or predisposition to hold the requirement of future steps of a problem in mind while working on any particular step of the problem.*

In addition to acknowledging the close relation of planning and problem solving, this definition furthermore implies a relation of planning to more basic cognitive abilities, such as working memory, specifically the mental representation of situational context ("the requirement of future steps") when working on any one part of a problem. The relation of planning to more basic cognitive constructs – such as working memory or response inhibition – has also been supported by recent research on planning ability (Gilhooly, 2005; Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). The executive functions[2] perspective of planning meshes well with the cognitive problem solving perspective described above, as executive functions can be defined as cognitive functions required for handling situations in which routine behavior is not sufficient (Kaiser, Mundt, & Weisbrod, 2005; Lezak, 1995) – which is a defining feature of problem situations (cf. Funke, 2003). Indeed, Zelazo, Carter, Reznik and Frye (1997) explicitly propose a problem-solving framework for studying executive functioning to investigate the integration of cognitive processes as part of a problem solving process. According to their view, "the function of executive function is to solve problems" (Zelazo et al., 1997, p. 219).

Although no exact and universal definition of executive functions exists, there is wide-spread agreement among researchers and practitioners that what lies at the heart of executive functioning is a set of cognitive abilities and processes that enable goal-directed yet adaptive behavior in changing environments (Jurado & Roselli, 2007; Kaiser et al. 2005; Royall, Lauterbach, Cummings, Reeve, Rummans, Kaufer, et al., 2002; Lezak,

---

[1] The term *planning ability* may either refer to the capacity for mental planning or – more commonly in an executive functions context – to performing planned behavior, including plan execution and monitoring.

[2] Depending on context *executive function* in singular may either refer to the global construct, also named *executive functioning*, or to one of its subcomponents, e.g., response inhibition or planning. Both meanings are encountered in the literature.

1995). This involves setting goals, making plans, initiating action, monitoring progress, inhibiting inappropriate responses, and modifying goals and plans as required by changes in the environment. Executive functions are involved in the organization of external behavior as well as internal thought processes. Because of their pervasive nature, they are relevant for effectively performing in many areas of life, be it in education, work settings or everyday situations.

The widely used definition by Lezak (1995, p. 650) includes four distinct components of executive functioning: (1) volition, (2) planning, (3) purposive action, and (4) effective performance. This definition emphasizes the functional aspects of executive control, i.e., the role that the underlying executive abilities play in relation to the environment (Burgess, Alderman, Forbes, Costello, Coates, Dawson, et al., 2006). Other models emphasize the processing aspects of the executive system, for example the central executive in Baddeley's (2002) model of working memory or the supervisory attentional system for behavior regulation proposed by Norman and Shallice (1986). A third approach is to use a psychometric definition based on factor analytic methods. To qualify as executive functions the factors identified in this manner should be distinguishable from domain-specific non-executive processes, such as perception or language. Studies of this type often identify several moderately correlated dimensions of executive functioning, such as working memory, attention, inhibition, set-shifting, rule finding, or planning and problem solving (e.g., Miyake et al., 2000; Royall et al., 2002).

Historically, executive functions have been closely associated with frontal cortex functioning, as the symptoms identified with an impairment in executive functions had first been observed in patients with traumatic brain damage to the frontal lobes (Luria, 1973; Stuss & Benson, 1986). Although modern brain imaging techniques have confirmed the important role of prefrontal brain regions for various executive functions, it has also become clear that other brain areas – including several subcortical structures – are also critically involved in supporting executive functions (e.g., Kaiser et al., 2005; Lezak, 1994; Royall et al., 2002). To the extent that the neurophysiology of executive functions is relevant for informing the cognitive and psychometric approach taken in this thesis, it will be selectively included in the corresponding manuscripts.[3]

Planning is generally considered a core component of executive functioning (Jurado & & Rosselli, 2007; Lezak, 1995). Indeed, Scholnick, Friedman and Wallner-

[3] As Miller et al. (1960, p. 196) remarked: "The procedure of looking back and forth between [psychology and neurophysiology] is not only ancient and honorable – it is always fun and occasionally useful."

Allen identify planning as the main role of the executive system: “The executive function acts as a master planner.” (Scholnick, Friedman & Wallner-Allen, 1997, p. 127). Similarly, in their problem solving framework of executive functioning Zelazo et al. (1997) consider planning a distinctive stage in the problem solving process facilitated by executive functions. A comprehensive review of the literature on executive functioning by Royall et al. (2002) also identified planning as a separable ability dimension. Additionally, practically all widespread test batteries for executive assessment contain at least one test explicitly aimed at measuring planning ability (cf. Jurado & Roselli, 2007).

Miyake et al. (2000) take a more differentiated psychometric view, suggesting that some executive abilities – such as planning – are complex higher-level functions that depend on other lower-level abilities. They empirically support this proposal using structural equation modeling, showing that performance in the Tower of Hanoi task can be partially explained by a lower-level inhibition factor. However, some of the results reported in the later manuscripts (Holt, Rodewald, Rentrop, Funke, Weisbrod, & Kaiser, 2011) suggest that performance in planning tasks may not be entirely reducible to other more basic cognitive abilities. A characteristic element of real-world planning tasks is to create structure in ill-structured, information-rich situations. As Burgess et al. (2006) argue, it is not obvious whether these complex demands can be adequately captured by a combination of well-structured and narrowly focused tasks such as the ones used in Miyake et al. (2000).

In summary, planning is a central element of executive functioning, although further research is required to determine its exact position in relation to other executive constructs and functional outcome. This in turn depends on the availability of suitable measures of planning ability, which will be discussed next.

1.4 Measuring planning ability

The most widely used paradigms for the neuropsychological assessment of planning ability (cf. Lezak, 1995; Friedman & Scholnick, 1997) are the Tower of London and its variants (Shallice, 1982) and maze tasks, in which participants have to find a way through a labyrinth using paper and pencil (Porteus, 1965). While both tasks are used in standardized neuropsychological assessment, the Tower of London probably is the dominant paradigm in recent research on planning ability in schizophrenia.

The Tower of London (TOL) task requires participants to transform an arrangement of three colored balls stacked on three pegs of different length from a given start state to a goal state, moving one ball at a time from one peg to another. The rules allow for only one ball to be moved at a time. Additionally the peg length limits the number of balls that may be stacked on it. Participants are usually instructed to mentally plan ahead in order to find the solution involving the minimal number of moves. The task is either presented physically or as a computer-based version. The original Tower of London was developed by Shallice (1982) and is closely related the classical Tower of Hanoi problem (Simon, 1975). Several variants of the task exist that are similar in nature and will henceforth be included under the label TOL unless stated otherwise.

Tests like the TOL are easy to administer and score, have reasonable psychometric reliability and show satisfactory convergent validity with respect to other tests of cognitive ability (e.g., Schnirman, Welsh, & Retzlaff, 1998; Tucha & Lange, 2004). These factors may partly explain their enduring popularity as research and assessment tools. However, their relation to real-life outcome measures, such as workplace performance or carrying out activities of daily living is often unclear or lacking (Burgess, Alderman, Volle, Benoit & Gilbert, 2009; Manchester, Priestley & Jackson, 2004; Shallice & Burgess, 1991). This is not entirely surprising, considering how little these standard neuropsychological tests resemble the challenges and complexities of everyday life. When the main purpose of neuropsychological testing is to assist with the diagnosis of brain pathology, this is not necessarily a problem. However, with a rising interest in ecological validity and predicting functional outcome, it becomes pertinent to ask to what extent existing neuropsychological test procedures possess ecological validity and how it can be increased (Burgess et al., 2006; Chaytor & Schmitter-Edgecombe, 2003; Chaytor, Schmitter-Edgecombe & Burr, 2006; Manchester et al., 2004). In the context of psychometric assessment, ecological validity can be defined as “the degree to which test performance corresponds to real-world performance” (Chaytor & Schmitter-Edgecombe, 2003, p. 182), i.e., the generalizability of test results beyond the test situation.

Burgess and colleagues (2006) have suggested several task demands that pose difficulties for patients with impaired executive functioning, yet are absent in many traditional neuropsychological tests. Among these are the ability to multi-task, work for comparatively long periods of time on a task without receiving feedback, and, more generally, to find a goal-oriented and structured approach to “ill structured” and complex tasks (Burgess et al., 2006; Goel, Grafman, Tajik, Gana & Danto, 1997). Taking the notion of ecological validity to its extreme, the Multiple Errands Test (MET; Shallice & Burgess, 1991) requires patients to carry out a set of errands in a real-world shopping center while being followed by an observer taking notes and scoring the patient’s effectiveness at doing so. This task was developed to address the perceived need for tests with high ecological validity, which comes at the expense of tying the test to local conditions, resulting in limited cross-site standardization. Moreover, conducting the test involves a comparatively high effort, even when using the simplified version (Knight et al., 2002) and it is difficult to create parallel versions for repeated testing. It appears that there is a gap between the highly standardized, reliable, scalable and economical testing that the TOL and its variants provide at one end of the spectrum and the high ecological validity and realism afforded by tests such as the MET at the other end.

1.5 Aims of the present thesis

Despite the initial evidence for the robustness and relevance of a planning deficit in schizophrenia, Reichenberg and Harvey conclude in their comprehensive review of cognitive impairments in schizophrenia that compared to other cognitive functions "dimensions of executive functioning such as planning and monitoring have received less research attention" (Reichenberg & Harvey, 2007, p. 842). In particular, despite robust evidence of a deficit in planning ability as measured by the TOL, there is sparse evidence of planning deficits in schizophrenia using other planning paradigms. It is therefore still largely an open question to what extent this deficit may be method-specific and also to what extent it generalizes to planning in everyday life. This issue is particularly acute considering that planning is a broad construct and the TOL is a very specific task with a strong visuo-spatial component that bears little similarity to most everyday planning situations (Burgess et al., 2006; Scholnick & Friedman, 1997). A likely reason for this paucity of research may be a lack of suitable alternative measurement paradigms that are reliable, construct-specific, economical to use, and ideally possess high ecologically validity. The first goal of the present thesis therefore was to develop a test of planning ability – named Plan-a-Day – that meets these requirements and fills the middle ground between traditional neuropsychological tests (e.g., the TOL) and real-world assessments (e.g., the MET). The research question associated with this test development was: *(1) Is the newly developed Plan-a-Day test a reliable and valid measure of planning ability in schizophrenia?*

A related second research question was how well planning ability as determined by neuropsychological assessment can predict everyday functioning of patients. Some authors are generally pessimistic about the ability of existing neuropsychological tests of executive functioning to predict everyday functioning. This is mainly due to the fact that the restricted and artificial testing situation lacks the unstructured nature and complexity that make everyday situations challenging for patients with executive deficits (e.g., Burgess et al., 2006; Lezak, 1995; Wilson, Evans, Emslie, Alderman, & Burgess, 1998). Studies combining standardized neuropsychological planning tasks and measures of functional outcome or functional capacity to assess the ecological validity of the standardized tests are still rare (e.g., Katz et al., 2007; Krabbendam, Vugt, Derix, & Jolles, 1999; Semkovska et al., 2004). It therefore was another goal of this thesis to investigate the ecological validity of standardized tests of planning ability, particularly that of the newly developed assessment method: *(2) Is planning ability – in particular as measured by the Plan-a-Day test – predictive of functional outcome?*

Third, assuming that planning is relevant for everyday functioning leads to the question whether training planning ability could in turn be used to enhance functional outcome in patients. Existing studies indicate that specific strategy-focused problem solving trainings may be effective in cognitive remediation in schizophrenia (e.g., Medalia, Revheim, & Casey, 2001; Krabbendam & Aleman, 2003) and specific planning trainings have been suggested as a way to improve everyday functioning in patients with schizophrenia (cf. Seter et al., 2011). As the Plan-a-Day concept has also been adapted as a computer-based training for cognitive remediation purposes (Holt & Funke, 2011), this lead to the third research question: *(3) Is a planning and problem solving training based on the Plan-a-Day concept effective in cognitive remediation?*

Finally, to investigate how disorder-specific the planning deficit is, it would be desirable to compare patients with schizophrenia to another psychiatric patient group. Additionally, multiple measures of planning should be employed in combination with a range of other neuropsychological ability measures. Combining these elements allows to address the question of the relative specificity of the deficit (Lewis, 2004) in a particularly comprehensive way, as confounds associated with patient/non-patient status of participants are eliminated and using multiple measures of planning ability enhances the generalizability of findings. At present, there is no published study that meets all of these criteria simultaneously. Another goal of this thesis therefore was to conduct a neuropsychological study of this type. The corresponding research question was: *(4) How specific is a planning deficit in schizophrenia?*

Beyond these explicit research questions, it was an implicit aim of this thesis to respond to Cronbach's call (1957, 1975) to join the "two disciplines" of experimental and differential psychology for a more complete and balanced approach. In the present case, this meant applying a process-focused cognitive perspective grounded in the psychology of problem solving to the psychometric task of creating a new test of planning ability. This combined perspective is evident in the principles guiding test construction as well as a later section reporting an analysis of cognitive processes during task performance using eye tracking and verbal protocols.

To address the research questions specified above, I first constructed a test of planning ability based on an errand scheduling paradigm – Plan-a-Day – as a tool for clinical assessment and research.[4] The aim was to create a test with a high degree of realism and corresponding ecological validity that is still psychometrically sound and economical to use. I then analyzed the reliability, construct validity and incremental validity of this new test in an empirical study with patients with schizophrenia (Manuscript 1). To supplement this psychometric approach, I investigated the cognitive processes involved in the task by reanalyzing data from two cognitive process tracing studies using eye tracking and verbal protocols (supplementary section). Second, I contributed to a randomized controlled trial investigating the effectiveness of a planning and problem solving training for patients with schizophrenia compared to a training of more basic neurocognitive abilities (Manuscript 2). Finally, I investigated whether there is empirical support for a specific planning deficit in schizophrenia by comparing the cognitive performance of patient groups with schizophrenia, depression and a healthy control group on a broad battery of tests of executive functioning with a focus on planning ability (Manuscript 3). The results of these studies were reported in three manuscripts:

1. Holt, D. V., Rodewald, K., Rentrop, M., Funke, J., Weisbrod, M., & Kaiser, S. (2011). The Plan-a-Day approach to measuring planning ability in patients with schizophrenia. *Journal of the International Neuropsychological Society*, 17, 327–35.
2. Rodewald, K., Rentrop, M., Holt, D. V., Roesch-Ely, D., Backenstrass, M., Funke, J., Weisbrod, M., & Kaiser, S. (2011). Planning and problem-solving training for patients with schizophrenia: a randomized controlled trial. *BMC Psychiatry*, 11, 73.
3. Holt, D. V., Wolf J., Funke, J., Weisbrod, M., & Kaiser, S. (submitted). Planning deficits in schizophrenia: Specificity, method independence and functional relevance. Submitted to *Schizophrenia Research*.

[4] When some activity was predominantly carried out by myself or to identify my personal opinion, I will use "I" in this manuscript. However, as is the case in most of science, many activities were carried out as team work or contain significant contribution from my collaborators. In these cases I will use "we" to signify this.

Additional published outcomes of this research program include several conference contributions, software for measuring and training planning ability, and a book chapter:

Holt, D.V. & Funke, J. (in press). *The Plan-a-Day test of planning ability* [Computer software]. Mödling: Schuhfried. Release scheduled for autumn 2012.

Funke, J. & Holt, D.V. (in press). Planen, Organisieren und Kontrollieren [Planning, organization and control]. In W. Sarges (Ed.), *Management Diagnostik*. Goettingen, Germany: Hogrefe.

Kaiser, S., Holt, D.V., Wolf, J., Rodewald, K., Rentrop, M, Funke, J., &Weisbrod, M. (2011). Specificity and relevance of planning deficits in schizophrenia. Talk presented at the *3rd European Conference on Schizophrenia Research*, 29 September – 1 October 2011, Berlin, Germany.

Holt, D.V. & Funke, J. (2011). PLAND – *Exekutive Funktionen: Planungs- und Handlungskompetenz* [PLAND – Executive Functions: Planning and action skills; computer software]. Mödling: Schuhfried. (Please refer to http://www.schuhfried.at for details.)

Rodewald, K., Rentrop, M., Holt, D., Roesch-Ely, D., Backenstrass, M., Funke, J., Weisbrod, M., et al. (2010). Targeting Planning and Problem Solving Versus Basic Cognition in Cognitive Remediation for Patients With Schizophrenia. *Schizophrenia Research (Proceedings)*, 117, 393–394.

Holt, D., Brüssow, S., & Funke, J. (2009). "What you see is what you say": On the convergent and predictive validity of eye movement recording and think-aloud protocols in a complex planning task. Poster presented a the *51. Tagung experimentell arbeitender Psychologen (TeaP) [51st Annual Conference of Experimental Psychologists]*, Jena, Germany, 29 March – 1 April 2009.

Brüssow, S., Holt, D., & Funke, J. (2008). Predicting eye movement behavior in a complex scheduling task using a cognitive process model derived from verbal protocols. Poster presented at the *9. Fachtagung der Gesellschaft für Kognitionswissenschaft (KogWis)* [9th Biannual Conference of the German Society for Cognitive Science], Dresden, Germany, 28 September – 1 October 2008.

2 The Plan-a-Day approach to measuring planning ability (Manuscript 1)[5]

To address the research questions stated above, the new Plan-a-Day test was constructed to measure planning ability reliably and in an ecologically valid manner. Plan-a-Day neither falls into the category of classical office-based tests – such as the TOL – nor into the category of real-life assessments – like the MET. It represents a computer-based simulation test, which provides a certain degree of realism and complexity while maintaining high control over the task (cf. Funke, 2001). Some computer-based simulation tests employ realistic three dimensional virtual reality environments (e.g., Kurtz, Baker, Pearlson & Astur, 2007; McGeorge, Phillips, Crawford, Garden, Della Sala, Milne, et al., 2001; Rand, Rukan, Weiss & Katz, 2009), whereas others consist of a schematic representation of a real-life situation, simplified with respect to visual presentation and possibilities for interaction (e.g., Craik & Bialystok, 2006; Larøi, Canlaire, Mourad & Van Der Linden, 2010). While the former approach is appealing in terms of approximating real life as closely as possible, the latter encourages a focus on task characteristics essential for assessment purposes as opposed to realistic but potentially incidental surface features. This has also been termed the *scaled worlds* approach to creating simulated environments (Gray, 2002).

Using a relatively complex computer-simulated task was relevant for addressing a particular challenge in the psychometric measurement of executive functions, that as has been succinctly summarized by Lezak (1995, p. 651):

> *A major obstacle to examining the executive functions is the paradoxical need to structure a situation in which patients can show whether and how well they can make structure for themselves.*

In essence, this statement implies an intrinsic contradiction between reliability and validity of tests of executive functions. In order to measure executive functions in an ecologically valid manner, the test situation should be novel and lack an obvious structure. However, to make the test objective and construct-specific, psychometric tasks are often narrowly defined with a clear structure and task goals imposed by instructions and test the administrator. This is proposed as one of the main reasons why traditional tests of executive functioning are not always predictive of real-world performance (Burgess et al., 2006; Burgess et al., 2009; Lezak, 1995). Furthermore, as handling novel, non-routine

[5] *This section accompanies the theoretical part of Manuscript 1, the empirical results are reported in the next section.*

situations is a defining element of executive functions (Kaiser et al., 2005; Lezak, 1995), it is difficult to increase psychometric reliability by administering the same task multiple times, as this would threaten the validity of the construct (Rabbitt, 1997). One way to answer this challenge is to combine a range of novel heterogeneous tasks into one test battery. This is the approach taken by the Behavioral Assessment of the Dysexecutive Syndrome (BADS; Wilson, Evans, Emslie, Alderman, & Burgess, 1996), which combines six different subtests and a clinical rating scale. The test-retest reliability of some subtests is only low to moderate but this is ameliorated by aggregating scores over different subtests. The downside of this approach is that the overall score is – although predictive of everyday executive difficulties – not particularly specific to any particular facet of executive functioning. Moreover, it incurs a high overhead in terms of test economy, as six different tasks need to be prepared and instructed. In the development of Plan-a-Day, we chose a different approach for dealing with this problem. The scaled worlds framework of Plan-a-Day, which requires participants to schedule a list of errands in a workplace setting, provides a flexible basis for creating a large number of fairly complex items that require participants to impose structure on the task. This makes it possible to administer a number of sufficiently different items thereby preventing simple strategy learning effects. At the same time, the design parameters that can be varied for item construction (see below) were deliberately restricted to facilitate systematic item construction and scoring. Whether this approach to balancing validity and reliability has been successful or not is the subject of the empirical part of Manuscript 1. If the test demonstrates psychometric reliability, construct specificity and ecological validity, then the design rationale is supported.

Finally, many existing neuropsychological planning tests, such as the Tower of London or maze tasks, are move planning tasks with a strong visuo-spatial component (e.g., Rushe et al., 1999). In contrast Plan-a-Day is a scheduling task that requires planning in the temporal domain, drawing mostly on verbally encoded information. As such, it aims to complement existing tests by covering another planning domain and a different mode of representation. Moreover, temporal scheduling problems are a common feature of everyday life, whereas most move planning problems do not have direct real-life equivalents.

The goals of developing Plan-a-Day can therefore be summarized as creating a new test of planning ability that allows the construction of a wide range of different test items, combining high ecological validity and a realistic task setting with good psychometric reliability. At the same time it should complement existing visuo-spatial planning tasks by providing a planning task that emphasizes the verbal-temporal domain.

2.1 The errand planning paradigm

Along with tower tasks and navigating mazes, errand planning tasks represent a frequently used paradigm in cognitive planning research (Scholnick et al., 1997). They usually require participants to devise a plan to carry out a given list of daily errands or plan a shopping trip given various constraints, e.g., time limits on when particular tasks can be carried out, see Figure 1 for an example. This task type emphasizes the role of planning as sequencing, in which “the planner’s task is to generate elements [of the plan], prioritize them, and sequence them effectively” (Scholnick et al., 1997, p. 129). The same authors argue that errands planning tasks are valuable in research because they relate to “meaningful, ordinary activities” relevant in everyday life (Scholnick & Friedman, 1993, p. 148). This may be an important reason why – besides the experimental use of daily errands tasks to study cognitive processes in planning (e.g., Hayes-Roth & Hayes-Roth, 1979; Dreher & Oerter, 1987; Gauvain & Rogoff, 1989) – there have been several proposals to turn this task type into a psychometric instrument for assessing planning ability.

You have just finished working out at the health club. It is 11:00 and you can plan the rest of your day as you like. However, you must pick up your car from the Maple Street parking garage by 5:30 and then head home. You’d also like to see a move today, if possible. Show times at both movie theaters are 1:00, 3:00, and 5:00. Both movies are on your “must see” list, but go to whichever one most conveniently fits into your plan. Your other errands are as follows:

> pick up medicine for your dot at the vet
> buy a fan belt for your refrigerator at the appliance store;
> check out two of the three luxury apartments;
> meet a friend for lunch at one of the restaurants;
> buy a toy for your dog at the pet store;
> pick up your watch at the watch repair;
> special order a book at the bookstore;
> buy fresh vegetables at the grocery;
> by a gardening magazine at the newsstand;
> go to the florist to send flowers to a friend in the hospital

Figure 1. Instructions for the errands planning task in Hayes-Roth & Hayes-Roth (1979, p. 277). The description is accompanied by a map showing the locations.

The Zoo Map task, included in the BADS (Wilson et a., 1996), requires participants to plan a visit to the zoo using a map, given a list of locations to visit and a set of rules to obey. The task has an obvious resemblance to everyday situations, despite some rather artificial constraints, e.g., that some foot paths may only be used once. A problem from a psychometric perspective is that this test consists only of one item without parallel forms, limiting its reliability and its use in repeated assessment. While this is not a major problem when using the test – as intended – as part of the complete BADS battery, it limits its usefulness as a specific standalone test of planning ability. The test-retest reliability of the Zoo Map test reported in Wilson et al. (1996) for an interval of 6 to 12 months was $r = .39$, $p = .03$. In our own research we observed a value of $r = .44$, $p < .01$, for a three-week interval (Holt et al., 2011). In terms of external validity, Katz et al. (2007) found that the Zoo Map correlated with instrumental activities of daily living, $r = .40$, $p < .05$, and Wykes et al. (2012) showed that an improvement in planning ability as measured by the Zoo Map was predictive of changes in work quality at the workplace in cognitive remediation therapy, beta = .48, $p = .02$.

Another more recent development also intended for neuropsychological assessment is the test for Action Organization and Daily Planning (HOTAP; Menzel-Begemann, 2010). It contains a subtest (HOTAP-C) that requires participants to form a sequence of action steps for several activities (e.g., making coffee, mowing the lawn) depicted as photographs. Participants simultaneously have to take into account additional constraints given in the instruction text. Like the Zoo Map, the HOTAP-C is a paper-based test using a single item with no parallel versions. After a two-year interval the test-retest reliability for the main summary score was satisfactory, $r = .69$, $p < .001$. Additionally, the correlations with performance on a daily planning task, $r = .39$, $p < .001$, and a post basket task, $r = .49$, $p < .001$, from a test of aptitude for clerical work (Marschner, 1981) emphasize the construct validity of this approach.

Similarly, the *Bogenhausener Planungstest* (BPT) is an earlier development by Stoltze (1991) also designed with the aim to assess planning ability in neurological patients. It requires scheduling a set of daily errands and activities (e.g., visiting the lost property office, meeting with relatives) using a set of cards representing the tasks and another set of cards describing the different locations and the paths between them. However, no systematic empirical investigations of the test's psychometric properties have been published to date.

More recently, planning tasks based on errand scheduling have also been implemented on the computer, allowing a more realistic dynamic interaction of task and participant and contributing to a more immersive test experience. The starting point for test presented in this thesis was a computer-based errand scheduling task developed by Funke

& Krüger (1993, 1995)[6]. This test – henceforth referred to as PAD95 – was inspired by post basket exercises commonly used in personnel selection (e.g., Jeserich, 1981) as well as the *Bogenhausener Planungstest* and aimed to address several shortcomings of existing planning tests and to leverage the use of computers in testing. The task is framed in workplace semantics and requires participants to schedule a list of work-related activities (e.g., picking up mail at the post office or checking inventory at a storage facility) while considering various constraints about when, where and for what duration the activities have to be carried out[7]. The difficulty of each problem is determined by the number of tasks to be scheduled and the number and interaction of constraints that need to be considered. In particular, the concept of Funke and Krüger emphasized realistic semantic embedding, improved psychometric reliability, detailed logging for process analysis, easy modification of items, and introduced a range of constraints and behavioral options to make the task more complex. For example, different priorities of errands need to be considered and a "taxi joker" is available which reduces the travel time for one errand to a third.

The reliability and validity of the test have been empirically assessed in a series of three studies involving students and managerial personnel, as reported in Funke and Krüger (1995). In one study involving N=104 university students, the test-retest reliability of the final score was shown to be moderate between r = .48 and r = .56. In a second study, a factor analysis of the tasks employed in an assessment center for personnel selection with N=78 participants revealed that the test formed a separate factor largely distinct from the other constructs assessed (e.g., teamwork, customer focus, decision making, leadership, etc.), suggesting a unique contribution towards explaining individual differences. Finally, a small experimental study comparing business managers to an unselected control group showed a significantly better performance of the manager group, $t(34) = 2.29$, $p < .05$, with a large effect size of Cohen's $d = 0.76$.[8] In sum, these findings indicate moderate levels of reliability and support the divergent and criterion validity of the PAD95 test. Although it

---

[6] The test developed by Funke & Krüger (1995) was also named Plan-A-Day. As it was the direct predecessor of the test presented in this thesis, we (DH and JF) decided to keep the name. Except for adopting the map and the generic kind of activities (e.g., "picking up mail at the post office") the test was completely redesigned. Details of similarities and differences will be discussed below. To avoid confusion, the test by Funke & Krüger (1995) will henceforth be referred to as PAD95.

[7] Due to the business semantics of Plan-a-Day, it would probably be more appropriate to speak of *tasks* that need to be carried out at different locations rather than *errands*. However, I will use the word errand to avoid confusions with the Plan-a-Day task itself and to emphasize that Plan-a-Day belongs to the class of errand scheduling tasks. Independent trials of Plan-a-Day involving the creation of different day plans will be referred to as problems or items.

[8] Effect size calculated from statistics reported in Funke and Krüger (1995).

has only been validated with students and managerial personnel the authors explicitly suggest an application of the test in neuropsychological assessment.

The validity of this approach to measuring planning ability is also supported by a similar more recent development, the Tour-Planner (Arling, 2006; Arling, Schellmann, & Spijkers, 2010), a computer-based visually enhanced version of a post basket task developed for use in occupational rehabilitation. The Tour-Planner requires planning a sightseeing tour through a city considering various constraints (e.g., opening times, entrance fees). The pen-and-paper version was predictive of aptitude for clerical work, $r = .56$, $p < .001$ (Arling, 2006), while the computer-based version also demonstrated external validity with respect to a work performance assessment in occupational rehabilitation, $r = .37$, $p < .01$ (Arling et al., 2010).

In sum, there is encouraging initial evidence for the validity and practical feasibility of the concept of computer-based errands scheduling tasks as proposed by Funke and Krüger (1995), although the psychometric properties – particularly reliability – should be investigated and improved further. Furthermore the test concept may need to be adapted for use with different target groups such as psychiatric patients.

2.2 The Plan-a-Day test

The test presented in this thesis is based on the original concept by Funke and Krüger (1995) with modifications to the basic task structure, a changed user interface, and a newly developed item set. As an introductory cover story, participants are asked to imagine that they work for a small company where they have to plan their daily activities. Information about the tasks to be carried out each day is presented in a task information area on the right side of the screen (see Figure 2), while the left side of the screen displays the different locations, distances between locations, and current position. The constraints that need to be considered for solving Plan-a-Day problems include earliest start time, latest finish time, location, duration of each errand as well as the distance between different locations. This format directly maps onto the Gantt chart method for planning, a commonly used project management technique (Field & Keller, 1998). The conceptual compatibility with a formal planning method supports the content validity of Plan-a-Day and is also useful for visualizing the structure of Plan-a-Day problems (for an example see Manuscript 1, Figure 2). When starting the task, participants are instructed to first plan their daily activities and then implement this plan by moving the symbol representing their current position. This is done by clicking on the corresponding locations on the map. The new location is then added to the “Schedule” display of the user interface (see Figure 2) with corresponding arrival and departure times. If it is too late to carry out a task at a location, a small information window with a corresponding message is displayed on screen. There is only one correct solution for each Plan-a-Day problem. When participants notice a mistake in their plan, they can undo previous moves using the “back”-button, which is accompanied by an information window confirming the undo operation as well as an acoustic signal. Participants are instructed to avoid using the “back”-button by planning ahead appropriately. Workplace semantics were chosen to increase the face validity of the task. This may be important for the acceptance of the test and, therefore, for motivation and engagement of participants. Furthermore, the semantics also help to activate pre-existing knowledge about approaching scheduling problems (cf. Blessing & Ross, 1996), which is desirable from a perspective of ecologically valid testing.

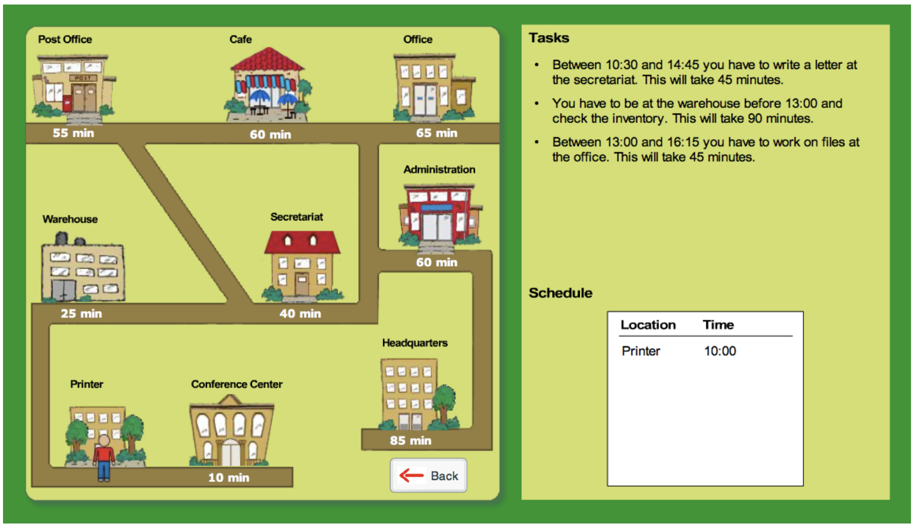

Figure 2. User interface of Plan-a-Day showing the initial state of a problem.

The test consists of two practice problems and eight assessment problems of increasing difficulty, requiring 20 to 30 minutes for completion. The operation of the program is explained and interactively demonstrated with the first practice problem using scripted instructions. Participants are then given the opportunity to further practice the operation of the program with the second practice problem. No time limit is set for the main assessment phase but the test program can be exited early after six of eight assessment problems should a participant seem overchallenged.

Compared to the version of the task developed by Funke & Krüger (1995), the present version has been modified and redesigned in several respects. Table 1 summarizes some of the major modifications. The motivation for these changes was to (a) further improve the psychometric qualities of the task, (b) adapt the task for a different target group (psychiatric patients), and (c) simplify the task structure to aid systematic item construction and analysis.

Table 1. Major modification of Plan-a-Day compared to Funke & Krüger (1995).

Test Feature	Funke & Krüger (1995)	Current Version
Errands per item	4–8	2–4
Maximum time per item	20 mins	3 mins
Items per test session	2 (out of 16)	8
Errand descriptions	separate window	on screen
Operation	keyboard	mouse
Taxi option	yes	no
Errand priorities	yes	no
Number of correct solutions	multiple	single

One of the main changes was the reduction of response options and constraints for the task (e.g., no taxi joker or priorities). This was done to simplify its computational structure and enhance its formal tractability (cf. Gray, 2002), which is important for rational item construction (see below) and for narrowing the scope of cognitive demands posed by the test. We aimed to reduce complexity to a level where executive demands are still present but the task becomes simple enough to allow construct-specific and reliable measurement of planning ability. To increase measurement reliability, we altered the test concept so that a larger number of short items rather than a smaller number of long items are administered. In order to simplify scoring we designed tasks to have only a single solution. Finally, we adapted the user interface, aiming to simplify operation (mouse instead of keyboard) and working memory demands (whole task in one window).

Formally stated, Plan-a-Day is a multiple constraint satisfaction problem that can be represented as a finite state machine. The correct solution is one particular path through the state space that needs to be found by the participant. On the cognitive level, Plan-a-Day requires (1) creating an appropriate mental representation of the problem which contains the essential features of the problem, (2) searching for a solution path in this mental model of the problem by applying various strategies and heuristics, before (3) implementing the solution in the task space and monitoring whether it is successful. As elaborated in chapter 4, which will discuss the cognitive processes underlying Plan-a-Day in detail, this sequence corresponds to a combination of what Newell & Simon (1972) described as the *planning method* and *heuristic search* for solving problems. In the taxonomy of planning processes proposed by Funke & Glodowski (1990) the abilities targeted by Plan-a-Day mainly lie in the domain of plan making, particularly the basic planning competencies of considering constraints and creating intermediate goals, i.e., making subplans. To a lesser extent it also covers the domain of plan execution, particularly monitoring of plan execution and taking corrective action if needed.

2.3 Item development and scoring

Initial item construction was based on a rational approach to test construction (cf. Moosbrugger, 2007), systematically varying the informational complexity of items. This was facilitated by the simplified task structure, which makes it easier to predict how changes in task parameters affect task difficulty. We operationalized informational complexity as the number of information elements (e.g., locations, start times, end times, durations, etc.) that need to be considered to find a solution, which in this task is necessarily correlated with the number of possible relations between these elements. While this may only be a rough approximation of complexity (cf. Funke, 2003), the operationalization is computationally simple, likely correlated with more elaborate measures of complexity, and easy to compare to measures of complexity used in other planning tasks, e.g., the number of moves required for a TOL problem.

Assessment problems were designed by systematically varying two factors influencing computational complexity: (1) the number of errands, with three levels ranging from two to four errands per problem, and (2) the number of different information elements that need to be considered to solve the problem. This factor also had three levels: problems that can be solved just by looking at start and end times (i.e., there is no overlap of the time frames for different errands), problems that also require considering errand durations due to time frame overlaps, and problems that additionally require errand durations and the time needed for reaching a particular location to resolve overlaps. This resulted in a 3 x 3 matrix for item construction, see Table 2. Easy items are displayed at the top left (few relevant information elements) and difficult items at the bottom right (many relevant information elements).

Table 2. Item design matrix.

Information Required	**Number of Errands**		
	2	**3**	**4**
A: Start and end times	Item 2A	Item 3A	Item 4A
B: Start and end times + durations	Item 2B	Item 4B	Item 4B
C: Start and end times + durations + distances	Item 2C	Item 3C	Item 4C

To test whether this scheme for constructing difficulty levels is valid, I used data from an additional study with a student sample not reported in the manuscripts. Assuming that each design factor is ordinally scaled implies an ordering of the difficulty of groups of items when the level of the other factor is held constant. If the test construction schema described above is valid this leads to the following hypotheses for the main performance variable solution time (*ST,* see below)[9]:

Hypothesis 1 – A larger number of errands increases item difficulty:

$$(ST_{2A} + ST_{2B} + ST_{2C}) < (ST_{3A} + ST_{3B} + ST_{3C}) < (ST_{4A} + ST_{4B} + ST_{4C})$$

Hypothesis 2 – More types of information required increase item difficulty:

$$(ST_{2A} + ST_{3A} + ST_{4A}) < (ST_{2B} + ST_{3B} + ST_{4B}) < (ST_{2C} + ST_{3C} + ST_{4C})$$

The sample consisted of $N = 98$ university students (67 female, mean age = 22.0 ± 3.81), who had completed the Plan-a-Day items listed in the design matrix, see Table 2. Two repeated-measures ANOVAs with the number of errands (levels: 2, 3, 4) and the information required (levels: A, B, C) as independent variables supported that both design factors significantly discriminated the items shown in the design matrix, $F_{\text{Errands}}(2, 194) = 227.51$, $p < .001$ (means for levels 2, 3, 4: 33.43 ± 27.47, 73.53 ± 56.46, 98.97 ± 50.71) and $F_{\text{Info}}(2, 194) = 22.47$, $p < .001$ (means for levels A, B, C: 60.85 ± 42.33, 64.41 ± 47.27, 80.66 ± 66.74). A detailed follow-up using directional linear difference contrasts confirmed that the levels of number of errands were ordered as hypothesized, $F_{2\text{ vs. }3}(1, 97) = 218.08$, $p < .001$, and $F_{3\text{ vs. }4}(1, 97) = 233.39$, $p < .001$. For information

[9] Subscripts refer to the corresponding items in Table 2.

required, the lower two levels were marginally different, $F_{A \text{ vs. } B}(1, 97) = 1.95$, $p = .08$, and the upper two levels were clearly different, $F_{B \text{ vs. } C}(1, 97) = 117.47$, $p < .001$. In sum, these analyses largely validate the construction schema described above. As we intended to keep the total test duration including instructions to less then 30 minutes, we removed item 2C from the final version of the test since we expected the more complex items with three or four errands to yield more diagnostically relevant information but wanted to keep two easy initial items.

Plan-a-Day was designed in such a manner that solution time captures most of the performance-relevant information. The program only proceeds to a new problem when the previous problem has been correctly solved, i.e., every problem is ultimately solved but the solution time may vary. As a side effect, this may improve participant motivation as every problem ends with a successful solution. This mechanism also means that planning mistakes incur a time penalty: undoing an incorrect move requires clicking the "back"-button and confirming the corresponding information message with another click before the move is undone. The intention of this design decision was to shift ability-related measurement variance from the error score to the time score, as accurate plans with fewer errors also results in faster solution times and vice versa. This has the added advantage of attenuating individual differences in speed-accuracy-tradeoff, as trying to increase speed at the expense of an increased number of errors in turn slows participants down again because of the corrections required to complete the item.

However, there are some disadvantages to using a time-based score as the primary measurement variable. Speed measures are more sensitive to interruptions and the effect of temporary inattention than purely power-based tests. Also, a personality disposition to be overly conscientious and focused on error prevention ("overplanning") may lead to solution times that underestimate true ability. The first aspect can be addressed by conducting the test under controlled conditions (e.g., avoiding interruptions), while the second aspect is harder to control but can at least be monitored by identifying and highlighting extreme values in the variable *planning ratio* (see below). We d not expect a strong impact of psychomotor retardation – which may be present in schizophrenia or depression – on the time-based scores in Plan-a-Day as compared to thinking time the motor execution is relatively short.

The scoring variables in Plan-a-Day are:

1. *solution time* (time from initial presentation to completion of problem): As described above, this is the main variable of Plan-a-Day.
2. *accuracy* (number of problems solved without error): The instructions of Plan-a-Day emphasize to construct complete, error-free plans first. Compared to solution time this variable is therefore expected to show comparatively little variance. However, in order to allow for comparison with other tests in which accuracy is the main measure, this variable can be scored. Alternatively, the average number of errors may be used as an accuracy measure, analogous to the number of excess moves in the TOL.
3. *planning time* (time from initial presentation to first move): Time spent deliberating before making the first move. This variable is mostly of interest as part of the planning ratio (see below).
4. *execution time* (time from first move to solution): An alternative to overall solution time, which implicitly measures planning quality as good plans will have fewer errors and therefore lead to shorter execution times.
5. *planning ratio* (proportion of planning time out of total solution time): As Manuscript 1 discusses, this variable appears to tap a construct different from solution time and may represent a tendency for reflected and strategic behavior.

2.4 Implementation

We chose to implement Plan-a-Day as a computer-based test, which provides several advantages (cf. Funke, 1998; Gray, 2002) such as:

- A more immersive and engaging experience due to the interactive nature of the task.
- Dynamic task elements, e.g., notifying the participant when it is too late to carry out a particular errand, followed by the opportunity to correct the error.
- Easy availability of response time measures, which form the main variable in Plan-a-Day.
- A rich process trace of user interactions recorded by the computer that can be used for process analysis.
- Automatic progression through the test under standardized conditions, improving test economy and objectivity.

Possible disadvantages include the effort to create a computer-based test, the requirement for computers to conduct the testing, and the role of computer experience as a possible confounding factor in test performance. While the first two are resource issues, the third point could affect the validity of measurement. However, perhaps somewhat surprisingly empirical research shows no definitive general effect of computer familiarity on performance in computer-based tests (Leeson, 2006). The increasingly important role of computers in everyday life may on the contrary even contribute towards ecological validity of such tests. However, the question of user interface design remains an important issue in computer-based testing (e.g., Booth, 1998), as cognitive demands, e.g., working memory load, can be affected by user interface design.

We therefore employed several principles of user interface design that are commonly used in computer-based tests. One of the most important principles was one-screen-only, i.e., that all task-relevant information is presented in just one window filling the whole screen. This avoids difficulties associated with navigating between multiple windows and the associated navigation-related working memory load. The only exception from this principle are small pop-up messages in the center of the screen displaying short messages (e.g., “You are too late to execute the task at this location.”), which can be closed with a single click. The second principle was a mouse-only operation to exclude keyboard and typing skills. The third principle was to reduce the number of possible user actions to a minimum. All Plan-a-Day tasks can be completed by just three mouse actions: (1) clicking on a location to move there, (2) clicking the “back”-button to undo the last move, and (3) confirming a message by clicking “OK”. In addition to these principles, we tried to make the visual aspects of the task as supportive as possible by using the full screen to enable a large display size, choosing fonts with good readability, and ensuring that important elements of the task were visually distinctive and easy to find on screen.

The test was programmed by the author in Adobe Flash 8.0 (http://www.adobe.com), using a modified model-view-controller paradigm with an event-driven message-passing mechanism. Adobe Flash was chosen as implementation platform, as it is a wide-spread web-based technology providing full interactivity. With small modifications the Plan-a-Day test could therefore also be delivered as a web-based test without the need to install additional software (“zero footprint”).

2.5 Future development

The Plan-a-Day test software with corresponding manual and standardized norms will be published commercially by Schuhfried (Austria) as part of the *Vienna Test System* (http://www.schuhfried.com), the release is scheduled for autumn 2012. This also involves a partial redesign of the user interface to match the look of the *Vienna Test System*, cf. Figure 3.

Holt, Daniel V. & Funke, Joachim (in press). *The Plan-a-Day test of planning ability* [Computer software]. Mödling: Schuhfried. Release scheduled for autumn 2012.

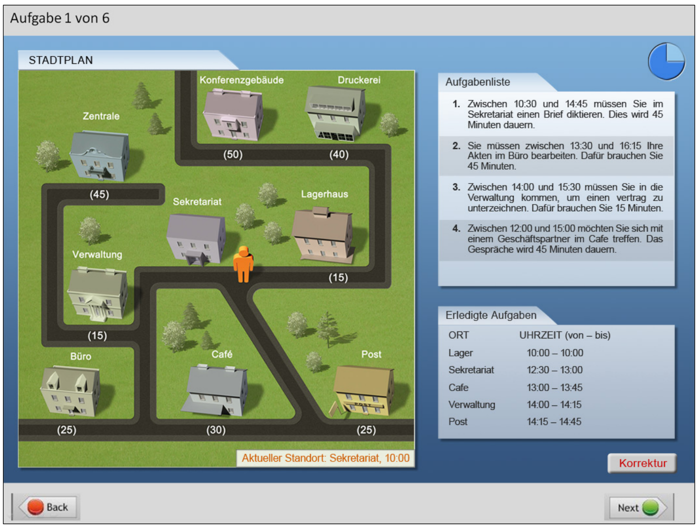

Figure 3. Design proposal for Plan-a-Day as part of the *Vienna Test System*. (Reproduced with permission of Schuhfried, Austria.)

3 Empirical validation of the Plan-a-Day approach in patients with schizophrenia (Manuscript 1)

Following the construction of the Plan-a-Day test we conducted an empirical study to investigate its reliability, construct validity and predictive validity with a sample of patients with schizophrenia. The results of this study are reported in detail in Manuscript 1, published in the *Journal of the International Neuropsychological Society* (Holt, Rodewald, Rentrop, Funke, Weisbrod, & Kaiser, 2011). As the theoretical basis for developing this test and the current state of research regarding the role of planning in schizophrenia have been described in the previous sections, I will now turn towards the empirical questions addressed in this study. These were: (1) whether Plan-a-Day shows good internal consistency and acceptable retest-reliability, (2) whether Plan-a-Day performance correlates with performance in other planning tests, indicating good construct validity and specificity, and (3) whether compared to standard neuropsychological tests Plan-a-Day can explain additional variance on a global level of functioning, demonstrating its ecological and incremental validity.

The study was conducted with a sample of *N*=80 patients with schizophrenia entering a rehabilitation program facilitating the return to a work environment. This patient group matched the design goals of Plan-a-Day particularly well, since – as outlined before – planning ability may be disproportionately affected in schizophrenia and the prospect of returning to a work environment places special importance on the ecological validity of a test of this ability. The high face validity of the simulated work place setting in Plan-a-Day should also render the test particularly acceptable for patients in this situation. This study was carried out as part of a project investigating the effects of a planning and problem solving training in cognitive remediation, which is reported in Manuscript 2. A neuro-psychological test battery including the Plan-a-Day test was administered twice with a four-week interval during which participants received cognitive ability training and inpatient work therapy. To introduce the Plan-a-Day test, the manuscript focuses on data from the first measurement. The global level of functioning was assessed using the Global Assessment of Functioning scale (GAF), a clinical rating scale that has been shown to possess good validity and reliability (Hilsenroth, Ackerman, Blagys, Baumann, Baity, Smith et al., 2000; Startup, Jackson, & Bendix, 2002). Furthermore, functional capacity was assessed based on patients performance in occupational therapy tasks using the Osnabrück Work Capabilities Profile (O-AFP; Wiedl & Uhlhorn, 2006). The neuropsychological test battery included a range of standard tests in the domain of executive functioning, as well as several tests that were specifically suited to assess the construct validity of Plan-a-Day, such as a TOL variant and the BADS Zoo Map to assess planning ability, as well as several forms of working memory assessment.

The results supported the psychometric qualities of Plan-a-Day with respect to all three questions. The internal consistency of Plan-a-Day's principal measure (solution time) was satisfactory (Cronbach's $\alpha = .78$) as was the test–retest reliability with an interval of four weeks ($r = .82$), comparing favorably with the other planning tests used in this study. As expected, Plan-a-Day solution time showed clear convergent validity with the other planning tests, the Tower of London and the Zoo Map. Discriminant validity was shown with respect to both verbal and spatial working memory, as well as crystallized intelligence as estimated by the MWT-B, supporting the construct-specificity of the test. In particular the discriminant validity of Plan-a-Day and the Corsi Block Tapping test of spatial working memory indicate that the design goal of creating a planning test with a low visuo-spatial component has been achieved. This may have been a critical point, as the map seems to suggest that spatial navigation is part of the task. However, in our view the map primarily serves as a visually attractive and structured way to present the verbal information, the distance estimates that can be obtained from the map are not precise enough to solve the task. Similarly, the near zero correlation of Plan-a-Day and the arithmetic test shows that it is not just a semantically framed test of mathematical ability, although simple calculations using various time-related errand properties (start time, end time, etc.) are part of the task. Plan-a-Day may be less characterized by mathematical complexity, but rather by the necessity to extract and integrate relevant information from the task environment.

Besides largely showing the expected pattern in terms of construct validity, Plan-a-Day also demonstrated incremental validity over the best alternative predictors used in this study in predicting global functioning as measured by the GAF scale. This is an encouraging result, as one aim of the development of Plan-a-Day was to move closer to measuring real-life functioning, for which the GAF is a first approximation. The finding indirectly supports the reasoning applied to test construction, such as using a realistic task setting to increase the transferability between the test situation and everyday life, as well designing a task with a realistic degree of complexity. These characteristics require participants to apply a considerable amount of strategic thinking, flexible adaptation of prior knowledge, and meta-cognition to develop and monitor a working strategy. No relation of Plan-a-Day to the O-AFP measure of functional capacity was observed at the first measurement occasion. Considering that no other cognitive test correlated with the O-AFP to a statistically significant extent, this may point at reliability or validity issues with this test, as discussed in Manuscript 1.

Considering secondary Plan-a-Day variables, the planning ratio emerged not only as a reasonably consistent overall measure, but also showed a characteristic pattern of relations to TOL accuracy, working memory, arithmetic ability and crystallized intelligence. We think that the planning ratio may indicate a well-planned and strategic approach to working on problems, possibly related to the construct of *cognitive reflection* (cf. Frederick, 2005). This would also explain why comparatively simple, less strategy-prone tasks (e.g., Trail Making Test A, Stroop) did not show a strong relation to this variable.

In summary, the results of this study show that Plan-a-Day is a valid and reliable instrument for measuring planning ability, with some indication that it also possesses incremental validity with regard to everyday functioning. Additionally, Plan-a-Day is easy to administer and score and we believe that the face validity afforded by the workplace semantics has a positive influence on participants' motivation and acceptance of the test. It appears that the Plan-a-Day approach shows some promise to fill the middle ground between traditional neuropsychological tests and real-world tasks, while offering interesting perspectives for both clinical application and research.

4 Cognitive process analysis using eye tracking and verbal protocols (supplementary)

To complement the psychometric validation of Plan-a-Day reported above from a cognitive perspective, this section provides additional data from two studies employing different cognitive process tracing methods (Brüssow et al., 2008; Holt et al., 2009). Please note that these data are included as supplementary materials and are not part of the manuscripts that constitute this publication-based thesis. I will select some central findings from these studies to illustrate how basic principles of human problem solving put forward by Newell & Simon (1972) can be applied to Plan-a-Day in order to understand the cognitive processes involved in the task, particularly with respect to the role of information reduction. As argued in the previous section, the requirement to select and organize information without external feedback may be a central feature of Plan-a-Day contributing to its ecological validity (cf. Burgess et al., 2006; Wilson et al., 1998). It therefore appeared worthwhile to investigate whether such processes of information reduction – as a cognitive response to informational complexity – are indeed evident in the process tracing data.

Studying mental planning poses a methodological problem: while responses recorded by the computer are useful for understanding the solution process, the phase before the first move – which is the time in which most of the actual planning occurs – does not generate any behavioral data. To study the details of cognitive processing, we therefore used two methods (eye tracking and verbal protocols) that indirectly provide information about cognitive processes without overt behavioral correlates. A core assumption underlying the use of eye movement recoding in cognitive research is the validity of the eye-mind assumption, i.e., that the overt focus of visual attention is related to current cognitive processes (Just & Carpenter, 1980). While this may not always be the case, e.g., due to covert attention shifts or the recall of information from memory (cf.

Viviani, 1990), eye movement recording has proven useful in the study of planning processes (e.g., Hodgson, Bajwa, Owen, & Kennard, 2000; Huddy, Hodgson, Kapasi, Mutsatsa, Harrison, Barnes, et al., 2007; Kaller, Rahm, Bolkenius & Unterrainer, 2009). Verbal protocols recorded from participants asked to "think aloud" while working on a task have also successfully been used for studying planning processes (e.g., Hayes-Roth & Hayes-Roth, 1979; Anzai & Simon, 1979; Gilhooly, Phillips, Logie, & Della Salla, 1999). However, despite representing an important source of data, verbalizing may affect task performance and verbalizations are also more intermittent and selective compared to the continuous stream of eye fixations (cf. Ericsson & Simon, 1993). We therefore triangulated the results of the eye tracking study with data from verbal protocols using an analogous data analysis procedure. Should the two data sources yield converging results this would increase our confidence in the validity of the analysis.

In the first study participants (N=42) worked on a range of different Plan-a-Day problems while their eye movements were recorded. The Plan-a-Day problems were visually simplified to facilitate eye movement recording, see Figure 4. In a complementary second study using the same stimulus materials participants (N=25) were asked to "think aloud" while working on the problems. Their verbalizations were digitally recorded and later transcribed and coded. The inter-rater reliability of the coding of verbal protocols was highly satisfactory with Cohen's kappa of .82. The analysis was based on Plan-a-Day problems 4A, 4B, and 4C, which are relatively information-rich and therefore particularly suitable for analyzing processes of information selection and reduction. The analysis was focused on the initial planning phase from the presentation of a problem up to the first behavioral response. For further details of the method please refer to the Appendix.

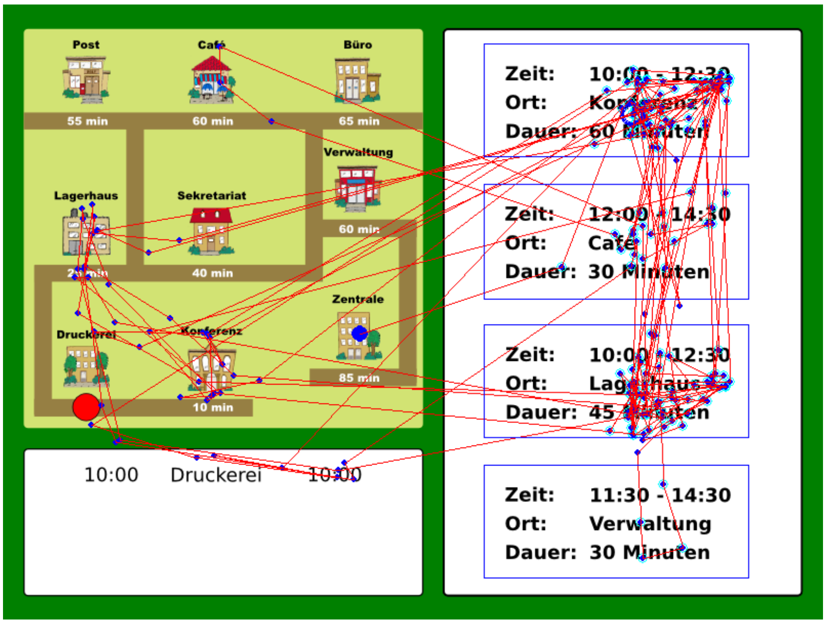

Figure 4. Plan-a-Day stimulus adapted for eye tracking with overlaid eye movement traces. Dots represent fixations and lines saccades, the boxes around errand descriptions identify the interest areas used for fixation analyses. In this instance the participant focused on the first and the third errand, deciding at an early stage not to consider the second and fourth errand any further.

Since Plan-a-Day was designed as a prototypical planning problem, we expected a pattern of information usage corresponding to the aptly named *planning method* described by Newell and Simon (1972). The planning method involves selecting a manageable part of the whole problem, forming an abstract representation of it that only includes information deemed relevant, mentally arranging a sequence of action steps in this planning space, and finally carrying out the plan while monitoring whether it works within the context of the whole problem. The planning method characteristically involves systematic information reduction, as only a part of the problem is selected to search for a solution, and even this part is represented only selectively by omitting irrelevant details. In his analysis of planning and problem solving using the Tower of Hanoi, Klauer (1993) named this type of information reduction *declarative simplification*, suggesting that people use systematic and elaborate processing strategies in problem solving but do so with simplified representations.

A strategy of systematic information reduction is furthermore facilitated by the fact that Plan-a-Day is a *set problem* (Simon & Newell, 1972), i.e., the correct solution is part of a – relatively small – set of possible options. Rather than randomly picking a possible solution candidate and testing whether it is correct, a rational problem solver may therefore use what Newell & Simon (1972) termed *heuristic search* to gradually eliminate unlikely options. Heuristic search can proceed according to different strategies, e.g., for a first move in Plan-a-Day dismissing errands for which the start time lies after the end time of another errand, or generally preferring tasks with early start times. In contrast to the *procedural simplification* (Klauer, 1993) inherent in heuristic decision making based on simple decision rules (e.g., Gigerenzer, 1999), heuristic search is an involved cognitive procedure where the search space is systematically narrowed down in several steps (Newell & Simon, 1972). This gradual process of information reduction with an ever narrower focus on the final solution implies a characteristic pattern of information usage in the process traces related to the order in which errands are presented and which errand is finally selected, as illustrated in Figure 5:

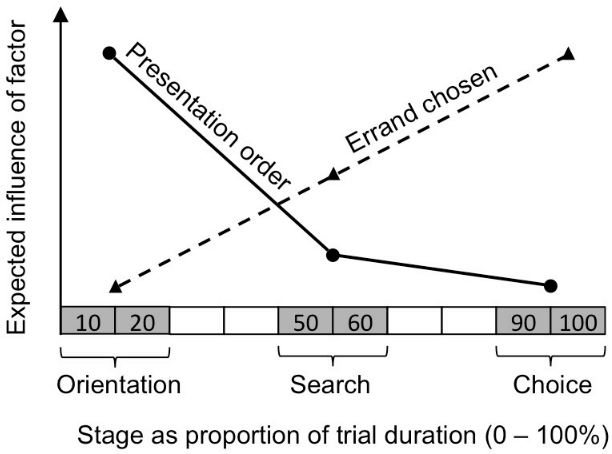

Figure 5. Schematic illustration of the expected influence of the presentation order of errands (top to bottom) and the errand finally chosen on explaining the distribution of eye fixations or verbalizations across errands.

1. *Orientation stage*: Participants read the errand descriptions to obtain an overview of the problem as a starting point before beginning to focus on a part of the problem. The number of references (eye fixations or verbalizations) related to each errand should therefore be determined mainly by the presentation order (top to bottom) of the errands.

2. *Intermediate search stage*[10]: As participants start to exclude options unlikely to represent the correct solution, the effect of initial presentation order diminishes compared

[10] For simplicity labeled "Search" in figures and statistical analyses.

to the orientation stage. A gradually increasing focus on the errand finally selected may start to become apparent at this stage, as partial selection has taken place already.

3. *Pre-choice stage*[11]: Immediately before choosing a response option the presentation order should be irrelevant, as the elimination of unlikely options has led to a very small set of options still under consideration. The effect of the errand finally chosen on the number of references (eye fixations or verbalizations) is now strong, as this errand should be part of this reduced set of options.

For statistical analysis, the stages were operationalized as a proportion of trial duration (stimulus presentation to first response) in order to compensate differences in absolute trial durations. To analyze the time course of information usage over these stages, time windows were chosen to be non-overlapping and equally spaced (orientation: 0-20% of trial duration, intermediate search: 40-60%, pre-choice: 80-100%), see Figure 5.

Earlier studies using eye movement recording with the Tower of London task (Hodgson et al., 2000; Huddy et al., 2007) demonstrated distinctive differences in eye fixation patterns between an initial orientation phase and subsequent elaborate problem solving phases. This distinction of stages corresponds to the orientation and intermediate search stages of the above model. Kaller et al. (2009) in turn found that the final gaze shift before making a choice in a Tower of London problem (corresponding to pre-choice stage of the above model) was a valid indicator of certain problem characteristics. Together, these findings support the distinction of stages proposed above and indicate that they have observable eye movement correlates.

As illustrated in Figure 6, results show an obvious effect of presentation order on the number of fixations received by each errand during the orientation stage, which disappeared at later stages. This was supported by a significant interaction effect of stage and presentation order in a two-way repeated-measures ANOVA, $F(6, 36) = 17.09$, $p < .001$, $\eta^2 = .43$, and by linear contrasts[12] showing the expected effects of presentation order at each stage, i.e., being statistically significant for the orientation stage but not for later stages: $F_{Orientation}(1,41) = 65.59$, $p < .001$, $\eta^2 = .62$; $F_{Search}(1,41) = 0.16$, $p = .69$, $\eta^2 = .00$; $F_{Choice}(1,41) < 0.01$, $p = .97$, $\eta^2 = .00$. An analogous analysis was carried out for the proportion of verbalizations referring to particular errands in the think aloud protocols, see Table 3, supporting the results from eye tracking. Verbalizations showed a significant

[11] For simplicity labeled “Choice” in figures and statistical analyses.

[12] A separate linear contrast was calculated for each stage, testing the statistical significance of the effect of presentation order as a linear trend (cf. Tabachnick & Fidell, 2007).

interaction effect for stage and presentation order, $F(6, 19) = 2.36$, $p = .02$, $\eta^2 = .10$, and corresponding results for linear contrasts of presentation order at each stage, even if the pattern was not as clear as for eye tracking: $F_{Orientation}(1,24) = 7.17$, $p = .01$, $\eta^2 = .23$; $F_{Search}(1,24) = 0.54$, $p = .47$, $\eta^2 = .02$; $F_{Choice}(1,24) = 2.11$, $p = .15$, $\eta^2 = .08$.

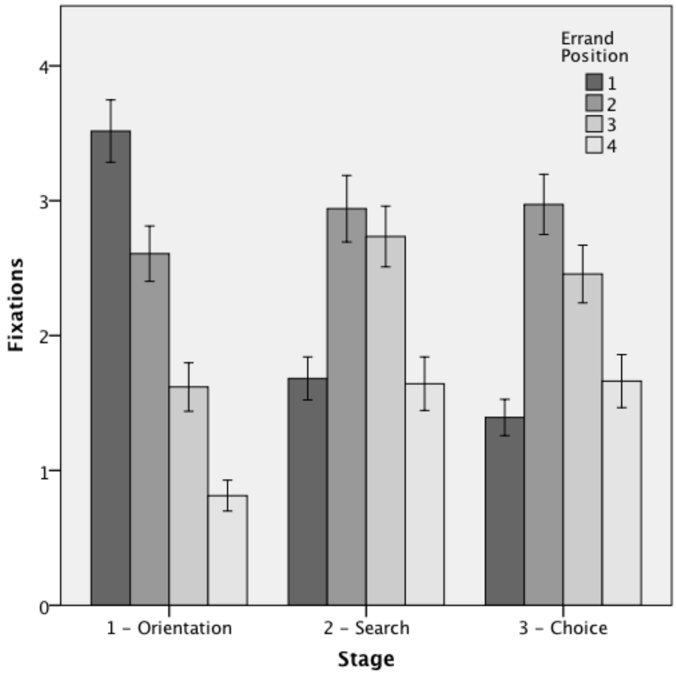

Figure 6. Average number of eye fixations on each errand information area during different stages prior to the first response. Errands are coded by presentation position (1 = top, 4 = bottom). Error bars indicate the standard error of the mean.

Table 3. Mean proportions of verbalizations referring to particular errands during different stages of working on the problem (standard error of the mean in parentheses).

		Stage		
		1 – Orientation	**2 – Search**	**3 – Choice**
Errand Position	1	.27 (.030)	.22 (.027)	.24 (.032)
	2	.31 (.026)	.26 (.032)	.21 (.033)
	3	.24 (.022)	.27 (.023)	.24 (.029)
	4	.18 (.014)	.25 (.029)	.31 (.030)
Errand Chosen	Other	.70 (.028)	.62 (.041)	.42 (.040)
	Chosen	.30 (.028)	.38 (.041)	.58 (.040)

As shown in Figure 7, the number of fixations on the errand chosen as a first move increased over time whereas the number of fixations on errands not chosen gradually decreased. A two-way repeated-measures ANOVA showed the expected interaction between number of fixations and stage, $F(2, 40) = 11.04$, $p < .001$, $\eta^2 = .21$. Furthermore, the follow up difference contrasts supported the increase of the proportion of fixations on the errand chosen from one stage to the next: $F_{Orientation\ vs.\ Search}(1, 41) = 8.58$, $p < .01$, $\eta^2 = .17$; $F_{Search\ vs.\ Choice}(1, 41) = 17.41$, $p < .001$, $\eta^2 = .30$. Again, the analogous analysis of verbalizations showed a similar pattern (see Table 3), namely an interaction of the proportion of verbalizations referring to the errand finally chosen and stage, $F(6,19) = 12.33$, $p < .001$, $\eta^2 = .34$. The follow up linear contrasts indicated that – as a slight variation of the eye tracking data – a clear focus on the finally selected response only became apparent for the late stage: $F_{Orientation\ vs.\ Search}(1,24) = 1.72$, $p = .20$, $\eta^2 = .07$; $F_{Search\ vs.\ Choice}(1, 24) = 21.72$, $p < .001$, $\eta^2 = .48$. This may be due to the fact that compared to the immediately available information from eye tracking, people verbalize their thoughts with a delay. Consequently the effect of this factor only becomes apparent somewhat later.

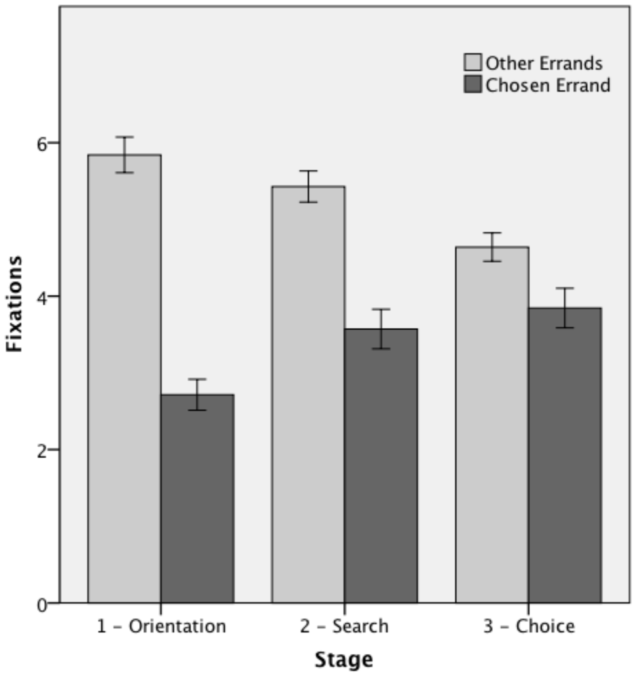

Figure 7. Number of eye fixations on the information area of the errand finally chosen and the errands discarded as response for the first move of a problem. Error bars indicate the standard error of the mean.

As a methodological side question, we also investigated the validity and reliability of the process tracing methods employed (Brüssow et al., 2008; Holt et al., 2009). The correlation of the number of eye fixations on a particular errand and the number of verbalizations referring to the same errand (in a by-stimulus analysis aggregated over all participants) was $r = .86$, $p < .01$, when considering the whole trial. When reducing the level of temporal aggregation by splitting each trial into ten intervals of equal duration the correlation between both methods was still $r = .56$, $p < .01$. This shows that aside from producing comparable results in the statistical tests reported above, both methods yield convergent evidence even on the level of raw data traces, mutually supporting their validity and reliability as indicators of the underlying cognitive processes. The predictive validity of both methods was further supported by the fact that the number of eye fixations on a particular errand allowed to predict the errand chosen as response with an accuracy of 71.0 % (chance level: 25%), while an analogous analysis for verbalization resulted in a predictive accuracy of 77.7 % (chance level: 25%).

Jointly these results support the assumption of a gradual process of information reduction as a characteristic feature of the Plan-a-Day task. Distinctive patterns of information usage were shown for three different stages of working on the task (orientation, search, pre-choice) with respect to influence of initial presentation order and the errand finally selected as response option. These stages match the stages identified in earlier eye tracking studies for the Tower of London (Hodgson et al., 2000; Huddy et al., 2007; Kaller et al., 2009). To efficiently perform this process of information reduction requires cognitive flexibility, as surface cues such as the initial presentation order or plausible but sometimes incorrect superficial errand-related cues (e.g., start times of task) need to be overcome in the course of finding the correct solution. The gradual search process should also benefit from a systematic and strategic approach and effective monitoring of options already considered. These cognitive demands implied by the information usage pattern observed in the process tracing data may contribute towards explaining the good ecological validity of Plan-a-Day, as shown in Manuscript 1. In general, the findings support the view that the requirement to select and organize information in order to create a plan is indeed a characteristic feature of Plan-a-Day.

5 The effects of a planning and problem-solving training in patients with schizophrenia (Manuscript 2)

While Manuscript 1 addressed the question of how planning deficits can be effectively measured, Manuscript 2 takes the next step from a clinical perspective and asks how and whether a planning and problem solving training is effective for improving cognitive impairments in patients with schizophrenia. The results of this study have been published in the journal *BMC Psychiatry* (Rodewald, Rentrop, Holt, Roesch-Ely, Backenstrass, Funke, Weisbrod, & Kaiser, 2011). There is converging evidence that cognitive remediation therapy is effective in improving cognitive performance in schizophrenia, particularly when combined with other means of rehabilitation (Bell, Fiszdon, Greig, Wexler, & Bryson, 2007; Cavallaro, Anselmetti, Poletti, Bechi, Ermoli, Cocchi, et al., 2009; McGurk, Twamley, Sitzer, McHugo, & Mueser, 2007). As several studies have shown that a change in executive functioning may be one of the best predictors of social and everyday functioning (Penadés, Catalán, Puig, Masana, Pujol, Navarro, et al., 2010; Reeder, Smedley, Butt, Bogner, & Wykes, 2006), this ability domain has become a focus area for cognitive rehabilitation efforts and corresponding research. Specifically, since planning and problem solving performance have been found to be related to functional outcome in schizophrenia (Aubin, Stip, Gélinas, Rainville, & Chapparo, 2009; Holt, et al., 2011; Seter et al., 2011), they may represent suitable target abilities for cognitive remediation. Training with a comparatively specific problem solving task, the Wisconsin Card Sorting Test, has been shown to transfer to a related cognitive task in patients with schizophrenia (Bellack, Weinhardt, Gold, & Gearon, 2001). Employing a more complex and interactive detective-scenario with problem-solving elements – the educational software "*Where in the USA is Carmen Sandiego?*" – even led to measurable improvements of skills of daily living as assessed by a structured interview (Medalia, Revheim, & Casey, 2001). Moreover, strategy-based trainings – such as sufficiently demanding problem solving trainings – generally seem to possesses a slight but measurable advantage over the training of more basic neurocognitive abilities (Krabbendam & Aleman, 2003). We therefore decided to follow up this lead in the present study by using a commercially distributed training version of Plan-a-Day (PLAN)[13], which represents a problem solving task with medium level of complexity and specificity compared to the tasks used by Bellack et al. (2001) and Medalia et al. (2001). PLAN has

[13] Until 2010 the software was distributed under the name PLAN as part oft the RehaCom package by Hasomed GmbH (Germany), since 2011 the revised edition is distributed as PLAND by Schuhfried (Austria).

been specifically adapted for cognitive remediation purposes with the aim to enhance activity planning and problem solving skills (Holt & Funke, 2011).

In contrast to the software used by Medalia et al. (2001) PLAN focuses on one particular type of problem solving, errands scheduling, which is common in everyday planning. Despite the different nature and scope of the task, PLAN nevertheless shares several characteristics with the task used by Medalia et al. and employs similar learning principles. It requires planning skills, organization, and deductive reasoning, and is motivating due to its simulated workplace context and by providing some form of personalization through adaptive progress. PLAN is based on a similar concept to the Plan-a-Day test, i.e., participants have to schedule different activities in different places using a map and time constraints. The training has three basic training modes, which require participants to complete tasks by priority, by minimizing travel distances, or by maximizing the number of errands completed. After familiarizing participants with the particular demands of each respective training mode, the training gradually progresses from simpler to more complex planning problems. The training is adaptive, i.e., the next level of difficulty is only introduced after two out of three tasks at a given level have been solved correctly. Figure 8 illustrates the user interface of the training version.

Figure 8. Screenshot of a feedback message during the practice phase of a training after making a mistake. (Reproduced with permission of Schuhfried, Austria.)

This study addressed the question whether targeting higher-level cognition – such as planning and problem solving ability – using the PLAN training provides an additional benefit over training basic cognitive functions (processing speed, attention, memory) in isolation. As Manuscript 1 has shown, the Plan-a-Day test delivered an increment in ecological validity that could not be reduced to more basic cognitive abilities. While these basic abilities certainly contribute to higher level cognitive functions such as planning (e.g., Miyake et al., 2000), the level of task complexity may produce emergent cognitive demands beyond the scope the basic component abilities, e.g., for selecting and organizing information, as has been argued in the previous sections. A comparable incremental effect of a training of complex cognition over basic cognitive abilities on functional outcome or its proxy measures may be therefore plausibly be expected.

The study was carried out as a randomized controlled trial comparing a planning and problem-solving training (Plan-a-Day) with a training of basic cognitive functions (processing speed, attention, memory). A group of patients with schizophrenia (N=77) participated in the study and received the training interventions in an inpatient rehabilitation setting parallel to a three-week course of inpatient work therapy.[14] The training interventions consisted of ten session of computer-based cognitive remediation therapy lasting 45 minutes each. Patients were randomly allocated to one of two treatment groups, and either practiced with PLAN or with three different tasks for training basic cognition (processing speed, attention, topological memory) presented on the same computer-based training platform as PLAN (RehaCom by Hasomed GmbH, Germany). During pre- and post-training assessments participants completed a comprehensive battery of neuropsychological tests while functional capacity was assessed using the O-AFP rating scales (Wiedl & Uhlhorn, 2006) in occupational therapy tasks.

Over the three-week training period both training groups improved in terms of functional capacity as assessed by the O-AFP, but no differential effects of the training methods were observed. The PLAN group showed a stronger improvement in planning ability than the basic cognition group as measured by the modified Plan-a-Day test and both groups showed an improvement on the TOL variant (Kohler & Beck, 2004) employed in this study. Additionally, the basic cognition group improved more on processing speed then the planning group.

The main hypothesis that a training of higher cognition would be more effective than a training of basic cognition was not confirmed, although both training groups

[14] The patient sample for this study was the same as in Manuscript 1.

achieved performance gains in the primary outcome measure. Several differences to the study of Medalia et al. (2001) may be relevant for interpreting this result. First, Medalia et al. (2001) used a multi-faceted problem-solving scenario involving a broad set of cognitive demands which was compared to a specific memory training. The present study, in contrast, employed a complex but still more focused problem solving task (errand scheduling) while the control group worked on a range of different basic cognitive abilities (processing speed, attention, topological memory). This choice of a broadly-based and therefore strong alternative treatment posed a higher hurdle for showing an incremental effect of the problem solving training compared to Medalia et al. (2001), i.e., it represented a more conservative study design. Second, the patients taking part in the present study were electively participating in a treatment program facilitating a return to work and had correspondingly mild cognitive impairments, while the sample in the study of Medalia et al. consisted of more severely affected inpatients. A tentative interpretation of the results from both studies may therefore be that – perhaps somewhat counter-intuitively – a training of higher-level cognition benefits patients with more severe cognitive impairments more than those with mild impairments. A possible explanation for this phenomenon is an interaction effect of higher-level cognitive strategies and basic cognitive abilities: The potential of patients' basic cognitive abilities (e.g., working memory) can only be applied effectively through the cognitive strategies they have available for approaching a task. Consequently, a low level of strategic competence becomes a performance bottleneck which may be masking the effect of the actual level of basic cognitive abilities. An improvement from low to moderate levels of strategic competence may therefore result in disproportionately strong performance gains by alleviating this bottleneck.

However, incremental effectiveness is not the only criterion for the usefulness of training methods. As noted above, the planning group improved more on the planning specific ability measure Plan-a-Day, the basic cognition group improved more on a measure of processing speed, and both groups showed improvements on the measure of functional capacity. While an unspecific treatment effect may be responsible for the improvement in functional capacity, this pattern is compatible with an alternative explanation: Both trainings may have led to similar improvements through different mechanisms. As a consequence, both training approaches may be suitable in practice, depending on the specific training needs of a patient. Moreover, the importance of maintaining motivation during cognitive rehabilitation, which is arguably supported by offering a choice of different training methods and including trainings set in meaningful contexts (cf. Medalia et al., 2001), makes PLAN an attractive option in a tool box of different training paradigms.

Overall, our results suggest that some cognitive deficits of patients with schizophrenia can be improved through a planning and problem solving training within three weeks, although no differential effects in comparison to a training of basic cognitive abilities could be established. In practical terms, PLAN can be viewed as a suitable component of a more broadly based training approach that includes multiple ability domains abilities and levels of task complexity. Indeed, this is the approach taken by the comprehensive computer-based training suites (Hasomed RehaCom and Schuhfried CogniPlus) of which PLAN and its successor PLAND are part.[15]

The author has created a revised edition of the PLAN software and manual in collaboration with the test publisher Schuhfried (Austria). As part of the Schuhfried *CogniPlus* training system this package has been released in 2011 under the name PLAND:

Holt, D.V. & Funke, J. (2011). *PLAND – Exekutive Funktionen: Planungs- und* Handlungskompetenz [PLAND – Executive Functions: Planning and action skills; Computer software]. Mödling, Austria: Schuhfried.

6 A planning deficit in schizophrenia: Specificity, method independence and functional relevance (Manuscript 3)

The third study of this thesis further broadens the scope of investigation from a specific new method of measuring and training planning ability towards the role of planning ability in schizophrenia in general. The results of this study have been presented as a conference contribution (Kaiser, et al., 2012) and have been submitted to the journal *Schizophrenia Research* for publication. As outlined in the introduction, schizophrenia affects a range of cognitive abilities (Heinrichs & Zakzanis, 1998) and it has been suggested that planning ability may be particularly affected (e.g., Morris et al., 1995; Frith, 1992). However, there is comparatively little evidence on the *relative specificity* (cf. Lewis, 2004) of a planning deficit in schizophrenia, i.e., whether there is a disproportionate deficit of planning ability when comparing the neuropsychological profile of patients with

[15] Please note that in Manuscript 2 PLAN is also referred to as Plan-a-Day, as this was the established name for the training version at the time. As the manuscript mostly focuses on the training version, we decided to distinguish the test version from the training version by explanations in the text when necessary.

schizophrenia to that of other patient groups. Alternatively, the planning deficit may not be domain-specific but simply an expression of an underlying more general cognitive deficit in schizophrenia (e.g., Dickinson, Ragland, Gold, & Gur, 2008). Figure 9 illustrates four basic possibilities for differences in neuropsychological profiles. We assume that the relative specificity hypothesis (see Figure 9D) applies to planning ability in schizophrenia.

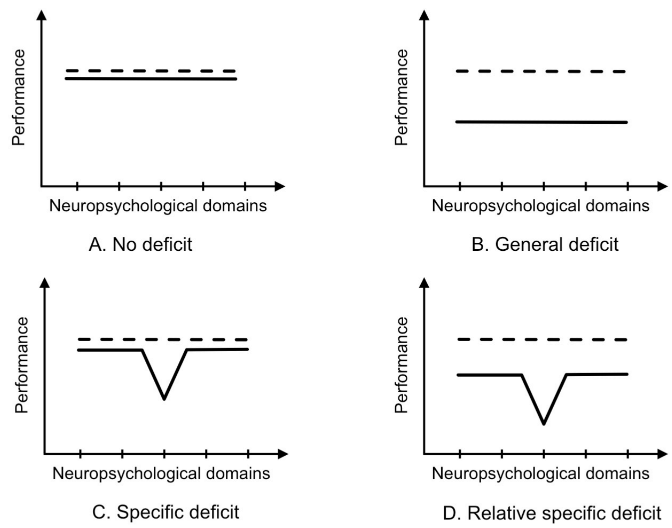

Figure 9. Schematic illustration of four basic deficit patterns when comparing the profile of two patients groups across several neuropsychological domains (attention, working memory, planning ability, etc.).

To investigate the relative specificity of a planning deficit, we compared patients with schizophrenia to another psychiatric patient group on multiple measures of planning together with a range of other neuropsychological ability measures. Combining these factors is a powerful approach to studying the specificity of a disorder-related cognitive deficit, that (a) eliminates confounds associated with patient/non-patient status of participants, (b) allows to compare the relative strength of impairment across different domains of functioning between patient groups, and (c) enhances the generalizability of findings in respect of planning ability as a construct in contrast to method-specific effects by employing multiple measures of planning ability. In doing so, this study extends earlier research comparing the neuropsychological profile of patients with schizophrenia and affective disorders (e.g., Goldberg, Gold, Greenberg, Griffin, 1993; Hill, Reilly, Harris, Rosen, Marvin, DeLeon, et al., 2009; Rogers, Kasai, Fukuda, Pletson, 2007) with a detailed coverage of planning ability. To our knowledge, this is the first study combining the features listed above to investigate the relative specificity of a planning deficit in schizophrenia.

As has been argued in the introduction of this thesis, we assume that the internal representation of context may be impaired in schizophrenia (Bustini et al., 1999; Cohen & Servan-Schreiber, 1992). We therefore expected a disproportionate planning deficit for the schizophrenia group compared to depression and a healthy control group, since the construction and maintenance of internal representations of context is particularly relevant for the planning and sequencing of complex actions (Cohen & Servan-Schreiber, 1992). While we expected a general deficit in executive functioning in major depression (Austin, Mitchell, & Goodwin, 2001), we did not expect planning ability to be disproportionately affected, as the representation of contextual information is not known to be specifically impaired in this disorder.

In addition to studying the relative specificity of a planning deficit in schizophrenia, we also investigated to what extent planning is predictive of functional outcome using three different measures. We expected to replicate the findings reported in Manuscript 1 with respect to the Global Assessment of Functioning and also added a self-rating scale of the dysexecutive syndrome as an exploratory outcome measure.[16] Furthermore, we recorded the number of days patients were medically certified as unfit to work during a six month period following cognitive assessment as a longitudinal element of the study. Besides being of interest to clinical practitioners this figure is a key statistical indicator in the German health care system. Being able to predict this variable would therefore be of practical relevance from a public health perspective.

The patients participating in this study (n=28 with schizophrenia and n=28 with depression) were recruited at the outpatient unit of a psychiatric hospital in addition to a control group of n=24 healthy volunteers. In order to show the specificity of a planning deficit in schizophrenia relative to other domains of cognitive functioning, we administered a battery of neuropsychological tests that covered five of the dimensions of executive functioning identified by Royall et al. (2002): planning, rule finding, working memory, attention and inhibition. Planning was assessed by three different tests covering different aspects of planning ability (Tower of London, Plan-a-Day and BADS Zoo Map).

[16] While it has been known that self-ratings of cognitive impairments in schizophrenia can be unreliable as strong impairments are associated with a lack of insight into the disorder (Medalia & Thysen, 2008), we exploratively included this measure since participants in this study were independently living outpatients and therefore possibly more reflected with regard to their condition.

We expected both patients groups to perform below the level of healthy control participants in general and anticipated a disproportionate deficit for planning ability in schizophrenia. Furthermore, we expected this deficit to replicate on different measures of planning ability. Finally, we investigated which neuropsychological variables were the best predictors of functional outcome.

The results largely conformed to expectations: Both patient groups showed a deficit in planning ability relative to the control group, but the deficit was significantly stronger for the schizophrenia group. The deficit was present across all three planning measures employed. On all other dimensions of executive functioning assessed in this study both patient groups scored comparably. Furthermore, together with working memory, planning ability was the best predictor of the Global Assessment of Functioning and of the number of days unfit to work recorded for both patient groups.

These findings extend existing research that showed replicable planning deficits in schizophrenia but did not include other patient groups, multiple measures of planning ability, and additional measures of executive functions. Planning ability indeed seems to be particularly affected in schizophrenia and this effect is not merely specific to a particular test of planning ability. Moreover, planning ability has once more been shown to be a good predictor of functional outcome, in this case also with respect to a longitudinally recorded behavioral outcome indicator, the number of days patients were medically certified as unfit to work.

7 Summary and conclusions

The research program presented in thesis set out with the goal to develop a new ecologically valid measure of planning ability and – using this tool – address several research questions with regard to the role of planning ability in schizophrenia. I started with the systematic development of a new test (Plan-a-Day) based on related earlier work. The test was then submitted to a comprehensive psychometric validation in a sample of patients with schizophrenia (Holt et al., 2011), which supported its reliability, construct validity and relation to functional outcome. In addition to the analyses reported in the manuscript, I tested the validity of the construction principles of the test with a student sample, which largely supported the design rationale for different difficulty levels. Furthermore, an additional reananalysis of data from two process tracing studies using eye tracking and verbal protocols (Holt et al., 2009; Brüssow et al., 2008) confirmed that the selection and reduction of information is a characteristic cognitive feature of the task. We then turned to the question whether the errand planning paradigm can also successfully be used as a training for cognitive remediation in schizophrenia (Rodewald et al., 2011). In a randomized controlled trial patients working with the PLAN planning and problem solving training showed an improvement of functional capacity comparable to that of another group who participated in a training of basic cognitive abilities. Finally, a study comparing different patient groups investigated the relative specificity of a planning deficit in schizophrenia and its predictive power in respect of functional outcome (Holt et al, submitted; Kaiser, Holt, Wolf, Rodewald, Rentrop, Funke, and Weisbrod, 2011). The results supported planning as a characteristic cognitive deficit in schizophrenia and its close relation to functional outcome.

Each one of the three main manuscripts contributes a novel element to the study of planning ability in schizophrenia that has not been present in the literature before. The study reported in Manuscript 1 was the first to investigate the psychometric properties of the Plan-a-Day approach to measuring planning ability in schizophrenia. More generally, it demonstrated the feasibility of adapting a daily errands scheduling task as a stand-alone psychometric test. The training study reported in Manuscript 2 was – to our knowledge – the first to compare the effect of a planning and problem solving training to that of a training of basic cognitive abilities with patients in a rehabilitation setting. Finally, the study presented in Manuscript 3 was the first to investigate the relative specificity of a planning deficit in schizophrenia compared to depression by employing multiple measures of planning ability as part of a comprehensive assessment of executive functions.

In the light of the evidence acquired in the empirical studies I will now revisit the guiding questions of this thesis.

(1) Is the newly developed Plan-a-Day test a reliable and valid measure of planning ability in schizophrenia? As the empirical data for a sample of patients with schizophrenia reported in Manuscript 1 show, the Plan-a-Day test possesses good reliability and its convergent validity and discriminant validity conformed to expectations. It correlated moderately with other measures of planning ability but not with spatial working memory or arithmetic ability. The test was easy to administer and score in practice, rendering its application feasible for both clinical and research use. Furthermore, from a research perspective Plan-a-Day represents an alternative to the widely used Tower of London in that it requires a different type of planning, verbal-temporal scheduling, instead of move planning. An additional study using a sample of university students also largely supported the validity of the design rationale for creating different difficulty levels

of the task. In addition to these psychometric investigations, a supplementary process analysis using eye tracking and verbal protocols provided further insight into the cognitive structure of information processing in Plan-a-Day. The process of creating structure in an informationally complex situations was hypothesized to be a key component of this test (Holt et al., 2011) and may be crucial for its ecological validity (cf. Burgess et al., 2006). The process tracing studies validated this assumption by showing a pattern of increasingly selective information processing compatible with general problem solving mechanisms described by Newell and Simon (1972).

(2) Is planning ability – in particular as measured by the Plan-a-Day test – predictive of functional outcome? Plan-a-Day demonstrated incremental validity over the next best predictor of functional outcome measured by the Global Assessment of Functioning as reported in Manuscripts 1 and 3. Together with working memory Plan-a-Day was also the best predictor of the number days that patients were medically attested to be unfit to work during six months period following neuropsychological assessment, as reported in Manuscript 3. These findings once more support the close relation of planning to functional outcome (Semkovska et al., 2004; Seter et al., 2011; Wykes et al., 2012) and underscore its clinical importance in neuropsychological assessment.

(3) Is a planning and problem solving training based on the Plan-a-Day concept effective in cognitive remediation? Manuscript 2 reported on a randomized control trial of a training of planning and problem solving ability using the PLAN training software. Findings showed that while the functional capacity of the group working with the problem solving training improved during the three-week training period, this improvement was comparable to a control group working with a training of basic cognitive abilities. Considering the importance of variety in training methods to maintain motivation and provide different training challenges renders PLAN a feasible option amongst others for designing a varied and stimulating training. This multiple-tool approach is evident in the basic concept of the comprehensive computer-based training suites (Hasomed RehaCom and Schuhfried CogniPlus) of which PLAN and its successor PLAND are part.

(4) How specific is a planning deficit in schizophrenia? Supporting the notion that a planning deficit may be a characteristic cognitive impairment in schizophrenia (e.g., Morris et al., 1995), the results reported in Manuscript 3 show that planning ability differentiated patients with schizophrenia not only from the control group but also from another patient group with depression. Other dimensions of executive functioning did not separate the two patient groups in this study. These findings support the central role of

planning ability in schizophrenia highlighting its relevance for both research and clinical practice. We agree with Bustini et al. (1999) that an impaired internal representation of context – as proposed by Cohen and Servan-Schreiber (1992) – may be a plausible and parsimonious explanation of such a deficit. This also allows to explain other characteristic cognitive impairments in schizophrenia, e.g., the disproportionate effect of increasing task complexity on performance observed in some studies (e.g., Hilti, et al., 2010; Morris et al., 1995; Marczewski et al., 2001). To establish the detailed mechanisms of such a specific processing deficit may be a promising direction for future research (cf. Cohen et al., 1999).

While some critical issues have been discussed in the previous sections and the corresponding manuscripts, I will now turn to some more general limitations and boundaries on the scope of the research presented together with corresponding suggestions for future research. Among these are, for example, whether a categorical distinction between different psychiatric disorders (e.g., schizophrenia and depression) is tenable, or whether a dimensional approach to psychopathology, e.g., based on the level of symptom severity in general or in several key areas, may be more appropriate also for explaining cognitive deficits (e.g., Greenwood, Morris, Sigmundsson, Landau, & Wykes, 2008; Stordal, Mykletun, Asbjørnsen, Egeland, Landrø, Roness, et al., 2005). As dimensional approaches to psychopathology are not yet particularly well established, we decided to employ a categorical system for these initial studies. Future studies could be designed to use a combined approach, like Stordal et al. (2005), by employing psychiatric rating scales that can be applied across diagnostic groups. Furthermore, the newly developed Plan-a-Day test presented in this thesis has so far only been thoroughly investigated with psychiatric patients. It is therefore an open question to what extent its psychometric properties will generalize to other populations and whether it will still measure the same construct, e.g., when the general cognitive ability of participants is very high. To address this question, we already have conducted another study with a sample of university students, who completed a test battery similar to that used in Manuscripts 1 and 3 but adapted for higher cognitive ability levels. A manuscript reporting the results of this investigation is currently prepared for publication.

On a more general level, the present thesis emphasizes the role of planning as a cognitive process and ability, largely leaving aside the details of how planning fits into a broader behavioral context of human self-regulation. As already laid out by Miller et al. (1960), planning plays a central role in the organization of behavior. The psychometric data on the ecological validity of Plan-a-Day support this connection: patients who scored higher on planning generally had a higher level of functioning in everyday life, where their planning skills are presumably used within a wider behavioral context. Indeed, this may be one of the most promising avenues to extend the present research. Just as we “drilled down” into the cognitive details of planning by carrying out psychometric analyses and detailed cognitive process studies, we may as well “move up” to higher level of behavioral organization to understand how planning processes are integrated into self-regulation in everyday life. Observational and naturalistic studies of planning (e.g., Alderman, Burgess, Knight, & Henman, 2003; Semkovska et al., 2004 ; Seter et al., 2011) have shown how investigations of this kind of research could be approached methodologically. This perspective links planning to other fields of research emphasizing the volitional and regulatory aspects of planning, such as the Theory of Planned Behavior (Ajzen, 1991), the concept of implementation intentions (Gollwitzer, 1999), or more generally in the study of prospective memory (e.g., Kliegel, McDaniel, & Einstein, 2000). Integrating this perspective into future research would contribute to a more holistic understanding of planning ability, particularly with a view towards explaining everyday deficits in patients with schizophrenia or other disorders involving executive impairments.

On a more methodological note, the investigation of some of the questions addressed in this thesis may be extended by applying structural equation modeling (SEM). SEM offers the attractive feature of providing correlations and mean structures of data that are in a statistical sense free of measurement error and highly construct-specific (Kline, 2011). However, the method does have several downsides which is why it was not chosen as the first step in this research program. Latent variable modeling techniques require large samples and multiple indicators for each construct assessed. Besides requiring significant amount of time and resources, this also is an ethical issues when working with patients. Should a study be inconclusive, the effort put in by patients may be lost. It therefore seems advisable to first establish empirical leads using smaller manifest variable studies and then follow them up using a latent variable approach. This also helps to address another potential problem in SEM studies, specification error (Kline, 2011). Using SEM effectively requires considerable prior knowledge to include the relevant variables in a study and adequately constrain the number and statistical parameters of the models being tested – which is facilitated by conducting prior studies. Finally, attractive as SEM is on a theoretical level, the manifest test scores and their relation to outcome indicators are relevant in individual neuropsychological assessment. So unless effects can also be shown on the manifest level, SEM may not be particularly helpful from a practical clinical perspective. For these reasons we decided to keep SEM approaches as a valuable second stage in a continuing research program on planning ability in schizophrenia.

To conclude, the research presented here supports that planning ability in general is a clinically relevant construct in schizophrenia with a clear relation to outcome measures. More specifically, results show that the new Plan-a-Day test is a reliable and valid measure of planning ability, which may be used in clinical application and basic research. A particular strength of Plan-a-Day from an applied point of view is its contribution towards predicting different measures of functional outcome. From a research perspective, it fills a niche in the range of available tests by providing a test of scheduling ability with a verbal-temporal emphasis not covered by other commonly used paradigms. Besides showing its value as a neuropsychological test, an adapted version of the Plan-a-Day concept was used in a training study which indicated that it performed comparably to other established training methods. Finally, a study comparing the neuropsychological profile of patients with schizophrenia to a group of patients with depression and a healthy control group supported the notion of a characteristic and practically relevant planning deficit in schizophrenia.

Overall, the research reported in this thesis illustrates that planning ability in schizophrenia is a rich topic of both theoretical and practical interest. It also is a field of research which benefits from combining both a psychometric and cognitive perspective to fully understand the phenomena under investigation. The present thesis aimed to address some pertinent issues in this field, but – as usual – many questions still remain open. I hope the studies presented here will encourage further research into the role of planning ability in schizophrenia, which indeed appears to represent a promising bridge construct between basic cognition and real-world functioning – just as Miller, Galanter and Pribram (1960) envisioned over fifty years ago.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179–211.

Alderman, N., Burgess, P. W., Knight, C., & Henman, C. (2003). Ecological validity of a simplified version of the multiple errands shopping test. *Journal of the International Neuropsychological Society*, 9, 31–44.

Anderson, J. R. & Douglass, S. (2001). Tower of Hanoi: evidence for the cost of goal retrieval *Journal of Experimental Psychology: Human Learning and Memory*, 27, 1331-1346.

Anderson, J. R., Albert, M. V., & Fincham, J. M. (2005). Tracing problem solving in real time: fMRI analysis of the subject-paced Tower of Hanoi. *Journal of Cognitive Neuroscience*, 17, 1261-1274.

Anderson, J. R. (1993). Problem solving and learning. *American Psychologist*, 48, 35–44.

Anzai, Y. & Simon, H. A. (1979). The theory of learning by doing. *Psychological Review*, 86, 124–140.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th ed., text rev.).* Washington, DC: Author.

Arling, V. & Spijkers, W. (2006). Planungskompetenz in der Beruflichen Rehabilitation - Entwicklung und Validierung eines strategischen Planspiels [Plannning ability in occupational rehabilitation – Development and validation of a strategic simulation game]. In Deutsche Rentenversicherung Bund. (Ed.), *15. Rehabilitations-wissenschaftliches Kolloquium: Rehabilitation und Arbeitswelt – Herausforderung und Strategien* (pp. 83–85). Berlin, Germany: Editor.

Arling, V. & Spijkers, W. (2010). Der computergestützte "Tour-Planer” [The computer-based "Tour-Planner”]. In Deutsche Rentenversicherung Bund. (Ed.), *19. Rehabilitations-wissenschaftliches Kolloquium: Qualität in der Rehabilitation – Management, Praxis, Forschung.* (pp. 144–145). Berlin, Germany: Editor.

Aubin, G., Stip, E., Gélinas, I., Rainville, C., & Chapparo, C. (2009). Daily activities, cognition and community functioning in persons with schizophrenia. *Schizophrenia Research*, 107, 313–318.

Austin, M.-P., Mitchell, P., & Goodwin, G. M. (2001). Cognitive deficits in depression:

Possible implications for functional neuropathology. *The British Journal of Psychiatry*, 178, 200–206.

Baddeley, A. D. (2002). Is working memory still working? *European Psychologist*, 7, 85–97.

Bell, M., Fiszdon, J., Greig, T., Wexler, B., & Bryson, G. (2007). Neurocognitive enhancement therapy with work therapy in schizophrenia: 6-month follow-up of neuropsychological performance. *Journal of rehabilitation research and development*, 44, 761–770.

Bellack, A. S., Weinhardt, L. S., Gold, J. M., & Gearon, J. S. (2001). Generalization of training effects in schizophrenia. *Schizophrenia Research*, 48, 255–62.

Blessing, S. B. & Ross, B. H. (1996). Content effects in problem categorization and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 792–810.

Booth, J. F. (1998). The user interface in computer-based selection and assessment: Applied and theoretical problematics of an evolving technology. *International Journal of Selection and Assessment*, 6, 61–81.

Burgess, P. W., Veitch, E., de Lacy Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, 38, 848–863.

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, UK: Psychology Press.

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L. M.-A., Dawson, D. R., et al. (2006). The case for the development and use of 'ecologically valid' measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society*, 12, 194–209.

Burgess, P. W., Alderman, N., Volle, E., Benoit, R. G., & Gilbert, S. J. (2009). Mesulam's frontal lobe mystery re-examined. *Restorative Neurology and Neuroscience*, 27, 493–506.

Bustini, M., Stratta, P., Daneluzzo, E., Pollice, R., Prosperini, P., & Rossi, A. (1999). Tower of Hanoi and WCST performance in schizophrenia: Problem-solving capacity and clinical correlates. *Journal of Psychiatric Research*, 33, 285–290.

Carroll, J. B. (1988). Individual differences in cognitive functioning. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology, Vol. 2: Learning and cognition (2nd ed.)* (pp. 813–862). Oxford, UK: Wiley.

Cavallaro, R., Anselmetti, S., Poletti, S., Bechi, M., Ermoli, E., Cocchi, F., et al. (2009). Computer-aided neurocognitive remediation as an enhancing strategy for schizophrenia rehabilitation. *Psychiatry Research*, 169, 191–196.

Chan, R. C. K., Chen, E. Y. H., Cheung, E. F. C., Chen, R. Y. L., & Cheung, H. K. (2004). Problem-solving ability in chronic schizophrenia: A comparison study of patients with traumatic brain injury. *European Archives of Psychiatry and Clinical Neuroscience*, 254, 236–241.

Chaytor, N. & Schmitter-Edgecombe, M. (2003). The Ecological Validity of Neuropsychological Tests: A Review of the Literature on Everyday Cognitive Skills. *Neuropsychology Review*, 13, 181–197.

Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. *Archives of Clinical Neuropsychology*, 21, 217–227.

Cohen, J. D. & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review*, 99, 45–77.

Cohen, J. D., Barch, D. M., Carter, C., & Servan-Schreiber, D. (1999). Context-processing deficits in schizophrenia: Converging evidence from three theoretically motivated cognitive tasks. *Journal of Abnormal Psychology*, 108, 120–133.

Craik, F. I. M. & Bialystok, E. (2006). Planning and Task Management in Older Adults: Cooking Breakfast. *Memory & Cognition*, 34, 1236–1249.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671–684.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, 30, 116–127.

Dickinson, D., Ragland, J. D., Gold, J. M., & Gur, R. C. (2008). General and specific cognitive deficits in schizophrenia: Goliath defeats David? *Biological Psychiatry*, 64, 823–827.

Dreher, M. & Oerter, R. (1987). Action planning competencies during adolescence and early adulthood. In S. L. Friedman, E. K. Scholnick, & R. R. Cocking (Eds.), *Blueprints for thinking: The role of planning in cognitive development.* (pp. 321–355). New York, NY: Cambridge University Press.

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

Field, M. & Keller, L. S. (1998). *Project Management.* Milton Keynes, UK: Open University.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 19, 25–42.

Frensch, P. A. & Funke, J. (Eds.). (1995). *Complex problem solving: The European perspective.* Hillsdale, NJ: Lawrence Erlbaum.

Frensch, P. A. & Funke, J. (1995). Definitions, traditions, and a general framework for understanding complex problem solving In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 3–25). Hillsdale, NJ: Lawrence Erlbaum.

Friedman, S. L. & Scholnick, E. K.Friedman, S. L. & Scholnick, E. K. (Eds.). (1997). *The developmental psychology of planning: Why, how, and when do we plan?* Mahwah, NJ: Lawrence Erlbaum.

Frith, C. D. (1987). The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychological Medicine*, 17, 631–648.

Frith, C. D. (1992). *The Cognitive Neuropsychology of Schizophrenia.* Hove, UK: Lawrence Erlbaum.

Fuller, R., Frith, C. D., & Jahanshahi, M. (2000). Reduced negative priming does indicate reduced cognitive inhibition in schizophrenia. *Cognitive Neuropsychiatry*, 5, 21–35.

Funke, J. (1998). Computer-based testing and training with scenarios from complex problem-solving research: Advantages and disadvantages. *International Journal of Selection and Assessment*, 6, 90–96.

Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking & Reasoning*, 7, 69-89.

Funke, J. (2003). *Problemlösendes Denken.* Stuttgart: Kohlhammer.

Funke, J. & Krüger, T. (1993). *"Plan-A-Day" (PAD*) [Computer software]. Bonn, Germany: University of Bonn, Department of Psychology.

Funke, J. & Krüger, T. (1995). "Plan-A-Day": Konzeption eines modifizierbaren Instruments zur Führungskräfte-Auswahl sowie erste empirische Befunde ["Plan-A-Day": Development of a modifiable instrument for selecting management personnel and first empirical results]. In J. Funke, T. Krüger, & A. Fritz (Eds.), *Neue Konzepte und Instrumente zur Planungsdiagnostik* (pp. 97–120). Bonn, Germany: Deutscher Psychologen Verlag.

Funke, J. & Glodowski, A.-S. (1990). Planen und Problemlösen: Überlegungen zur neuropsychologischen Diagnostik von Basiskompetenzen beim Planen. *Zeitschrift für Neuropsychologie*, 1, 139–148.

Funke, J. & Holt, D.V. (in press). Planen, Organisieren und Kontrollieren [Planning, organization and control]. In W. Sarges (Ed.), *Management Diagnostik*. Goettingen, Germany: Hogrefe.

Gauvain, M. & Rogoff, B. (1989). Collaborative problem solving and children's planning skills. *Developmental Psychology*, 25, 139–151.

Gazzaniga, M. S. (2000). *Cognitive neuroscience: A reader.* Oxford, UK: Blackwell.

Gigerenzer, G. & Todd, P. M. (1999). *Simple heuristics that make us smart.* New York, NY: Oxford University Press.

Gilhooly, K. J. (2005). Working memory and planning In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 71–88). Hove, UK: Psychology Press.

Gilhooly, K. J., Phillips, L. H., Wynn, V., Logie, R. H., & Della Sala, S. (1999). Planning processes and age in the five-disc Tower of London task. *Thinking & Reasoning*, 5, 339–361.

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

Goldberg, T. E., Gold, J. M., Greenberg, R., & Griffin, S. (1993). Contrasts between patients with affective disorders and patients with schizophrenia on a neuropsychological test battery. *The American Journal of Psychiatry*, 150, 1355–1362.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, 54, 493–503.

Gray, W. D. (2002). Simulated task environments: The role of high-fidelity simulations, scaled worlds, synthetic environments, and laboratory tasks in basic and applied cognitive research. *Cognitive Science Quarterly*, 2, 205–227.

Green, M. F. (1996). What are the functional consequences of neurocognitive deficits in schizophrenia? *The American Journal of Psychiatry*, 153, 321–330.

Green, M. F., Kern, R. S., Braff, D. L., & Mintz, J. (2000). Neurocognitive deficits and functional outcome in schizophrenia: Are we measuring the "right stuff"? *Schizophrenia Bulletin*, 26, 119–136.

Greenwood, K. E., Morris, R., Sigmundsson, T., Landau, S., & Wykes, T. (2008). Executive functioning in schizophrenia and the relationship with symptom profile and chronicity. *Journal of the International Neuropsychological Society*, 14, 782–792.

Greenwood, K. E., Wykes, T., Sigmundsson, T., Landau, S., & Morris, R. G. (2011). Tower of London versus real life analogue planning in schizophrenia with disorganization and psychomotor poverty symptoms. *Journal of the International Neuropsychological Society*, 17, 474–484.

Hayes-Roth, B. & Hayes-Roth, F. (1979). A cognitive model of planning. *Cognitive science*, 3, 275–310.

Heinrichs, R. W. & Zakzanis, K. K. (1998). Neurocognitive deficit in schizophrenia: A quantitative review of the evidence. *Neuropsychology*, 12, 426–45.

Hill, S. K., Reilly, J. L., Harris, M. S. H., Rosen, C., Marvin, R. W., DeLeon, O., et al. (2009). A comparison of neuropsychological dysfunction in first-episode psychosis patients with unipolar depression, bipolar disorder, and schizophrenia. *Schizophrenia Research*, 113, 167–175.

Hilsenroth, M. J., Ackerman, S. J., Blagys, M. D., Baumann, B. D., Baity, M. R., Smith, S. R., et al. (2000). Reliability and validity of DSM-IV axis V. *The American Journal* *of Psychiatry*, 157, 1858–1863.

Hilti, C. C., Delko, T., Orosz, A. T., Thomann, K., Ludewig, S., Geyer, M. A., et al. (2010). Sustained attention and planning deficits but intact attentional set-shifting in neuroleptic-naïve first-episode schizophrenia patients. *Neuropsychobiology*, 61, 79–86.

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 Neuroscience*, 12, 894–907.

Holt, D., Brüssow, S., & Funke, J. (2009). "What you see is what you say": On the convergent and predictive validity of eye movement recording and think-aloud protocols in a complex planning task. Poster presented a the *51. Tagung experimentell arbeitender Psychologen (TeaP)* [51st Annual Conference of Experimental Psychologists], Jena, Germany, 29 March – 1 April 2009.

Holt, D.V. & Funke, J. (2011). PLAND – *Exekutive Funktionen: Planungs- und Handlungskompetenz* [PLAND – Executive Functions: Planning and action skills; computer software]. Mödling, Austria: Schuhfried.

Holt, D.V. & Funke, J. (in press). *The Plan-a-Day test of planning ability* [Computer software]. Mödling, Austria: Schuhfried.

Holt, D. V., Rodewald, K., Rentrop, M., Funke, J., Weisbrod, M., & Kaiser, S. (2011). The Plan-a-Day approach to measuring planning ability in patients with schizophrenia. *Journal of the International Neuropsychological Society*, 17, 327–35.

Holt, D. V., Wolf J., Funke, J., Weisbrod, M., & Kaiser, S. (submitted). Planning deficits in schizophrenia: Specificity, method independence and functional relevance.

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

Hutton, S. B., Puri, B. K., Duncan, L.-J., Robbins, T. W., Barnes, T. R. E., & Joyce, E. M. (1998). Executive function in first-episode schizophrenia. *Psychological Medicine*, 28, 463–473.

Jeserich, W. (1981). *Mitarbeiter auswählen und fördern: Assessment-Center-Verfahren* [Selecting and developing personnel: Assessment center methods]. München, Germany: Hanser.

Jogems-Kosterman, B. J. M., Hulstijn, W., Wezenberg, E., & van Hoof, J. J. M. (2006). Movement planning deficits in schizophrenia: Failure to inhibit automatic response tendencies. *Cognitive Neuropsychiatry*, 11, 47–64.

Johnstone, E. C. & Frith, C. D. (1996). Validation of three dimensions of schizophrenic symptoms in a large unselected sample of patients. *Psychological Medicine*, 26, 669–679.

Jurado, M. B. & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, 17, 213–233.

Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329–354.

Kaiser, S., Holt, D.V., Wolf, J., Rodewald, K., Rentrop, M, Funke, J., & Weisbrod, M. (2011). Specificity and relevance of planning deficits in schizophrenia. Talk presented at the *3rd European Conference on Schizophrenia Research*, 29 September – 1 October 2011, Berlin, Germany.

Kaiser, S., Mundt, C., & Weisbrod, M. (2005). Exekutive Kontrollfunktionen und Neuropsychiatrische Erkrankungen – Perspektiven für Forschung und Klinik [Executive control functions and neuropsychiatric disoderds – Perspectives for research and clinical practice]. *Fortschritte der Neurologie, Psychiatrie*, 73, 438–450.

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.

Katz, N., Tadmor, I., Felzen, B., & Hartman-Maeir, A. (2007). The Behavioural Assessment of the Dysexecutive Syndrome (BADS) in schizophrenia and its relation to functional outcomes. *Neuropsychological Rehabilitation*, 17, 192–205.

Klauer, K. C. (1993). *Belastung und Entlastung beim Problemlösen: Eine Theorie des deklarativen Vereinfachens* [Cognitive load in problem solving: A theory of declarative simplification]. Göttingen, Germany: Hogrefe.

Kliegel, M., McDaniel, M. A., & Einstein, G. O. (2000). Plan formation, retention, and execution in prospective memory: A new approach and age-related effects. *Memory*

& Cognition, 28, 1041–1049.

Kline, R. B. (2011). *Principles and practice of structural equation modeling (3rd ed.).* New York, NY: Guilford Press.

Knight, C., Alderman, N., & Burgess, P. W. (2002). Development of a simplified version of the multiple errands test for use in hospital settings. *Neuropsychological Rehabilitation*, 12, 231–255.

Kohler, J. & Beck, U. (2004). *Planungstest* [Planning Test; computer software]. Konstanz, Germany: Beck und Kohler.

Krabbendam, L., de Vugt, M. E., Derix, M. M. A., & Jolles, J. (1999). The Behavioural Assessment of the Dysexecutive Syndrome as a tool to assess executive functions in schizophrenia. *Clinical Neuropsychologist*, 13, 370–375.

Krabbendam, L. & Aleman, A. (2003). Cognitive rehabilitation in schizophrenia: A quantitative analysis of controlled studies. *Psychopharmacology*, 169, 376–382.

Kurtz, M. M., Baker, E., Pearlson, G. D., & Astur, R. S. (2007). A virtual reality apartment as a measure of medication management skills in patients with schizophrenia: A pilot study. *Schizophrenia Bulletin*, 33, 1162–1170.

Leeson, H. V. (2006). The mode effect: A literature review of human and technological issues in computerized testing. *International Journal of Testing*, 6, 1–24.

Lewis, R. (2004). Should cognitive deficit be a diagnostic criterion for schizophrenia? *Journal of Psychiatry & Neuroscience*, 29, 102–113.

Lezak, M. D. (1995). *Neuropsychological assessment (3rd ed.).* Oxford, UK: Oxford University Press.

Lezak, M. D. (1994). Domains of behavior from a neuropsychological perspective: The whole story. In W. D. Spaulding (Ed.), *Integrative views of motivation, cognition, and emotion* (pp. 23–55). Lincoln, NE: University of Nebraska Press.

Luria, A. R. (1973). The frontal lobes and the regulation of behavior. In K. H. Pribram & A. R. Luria (Eds.), *Psychophysiology of the frontal lobes* (pp. 3–26). New York, NY: Academic Press.

Manchester, D., Priestley, N., & Jackson, H. (2004). The assessment of executive functions: coming out of the office. *Brain Injury*, 18, 1067–1081.

Marczewski, P., Van der Linden, M., & Larøi, F. (2001). Further investigation of the supervisory attentional system in schizophrenia: Planning, inhibition, and rule abstraction. *Cognitive Neuropsychiatry*, 6, 175–192.

Marschner, G. (1981). *Büro-Test [Office Test].* Göttingen, Germany: Hogrefe.

McGeorge, P., Phillips, L. H., Crawford, J. R., Garden, S. E., Della Sala, S., Milne, A. B., et al. (2001). Using virtual environments in the assessment of executive dysfunction. *Presence: Teleoperators and Virtual Environments*, 10, 375–383.

McGrath, J., Saha, S., Chant, D., & Welham, J. (2008). Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiologic Reviews*, 30, 67–76.

McGurk, S. R., Twamley, E. W., Sitzer, D. I., McHugo, G. J., & Mueser, K. T. (2007). A meta-analysis of cognitive remediation in schizophrenia. *American journal of psychiatry*, 164, 1791–1802.

Medalia, A., Revheim, N., & Casey, M. (2001). The remediation of problem-solving skills in schizophrenia. *Schizophrenia Bulletin*, 27, 259–267.

Medalia, A. & Thysen, J. (2008). Insight into neurocognitive dysfunction in schizophrenia. *Schizophrenia Bulletin*, 34, 1221–1230.

Menzel-Begemann, A. (2010). *HOTAP Handlungsorganisation und Tagesplanung [Action Organization and Daily Planning].* Göttingen, Germany: Hogrefe.

Mesulam, M.-M. (2002). The human frontal lobes: Transcending the default mode through contingent encoding. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function.* (pp. 8–30). New York, NY US: Oxford University Press.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior.* New York, NY: Holt, Rinehart, & Winston.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49–100.

Morris, R. G., Rushe, T., Woodruffe, P. W. R., & Murray, R. M. (1995). Problem solving in schizophrenia: A specific deficit in planning ability. *Schizophrenia Research*, 14, 235–246.

Nathaniel-James, D. A., Brown, R., & Ron, M. A. (1996). Memory impairment in schizophrenia: Its' relationship to executive function. *Schizophrenia Research*, 21, 85–96.

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

Norman, D. & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. Davidson, G. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation, Vol. 4.* (pp. 1–18). New York: Plenum Press.

Osvath, M. & Osvath, H. (2008). Chimpanzee (Pan troglodytes) and orangutan (Pongo abelii) forethought: Self-control and pre-experience in the face of future tool use. *Animal Cognition*, 11, 661–674.

Pantelis, C., Barnes, T. R., Nelson, H. E., Tanner, S., Weatherley, L., Owen, A. M., et al. (1997). Frontal-striatal cognitive deficits in patients with chronic schizophrenia. *Brain*, 120, 1823–1843.

Penadés, R., Catalán, R., Puig, O., Masana, G., Pujol, N., Navarro, V., et al. (2010). Executive function needs to be targeted to improve social functioning with Cognitive Remediation Therapy (CRT) in schizophrenia. *Psychiatry Research*, 177, 41–45.

Porteus, S. D. (1965). *Porteus Maze Tests: Fifty years' application.* Oxford England: Pacific Books.

Rabbitt, P. (1997). Introduction: Methodologies and models in the study of executive function. In P. Rabbitt (Ed.), *Methodologies and models in the study of executive function* (pp. 1–38). Hove, UK: Psychology Press.

Raby, C. R., Alexis, D. M., Dickinson, A., & Clayton, N. S. (2007). Planning for the future by western scrub-jays. *Nature*, 445, 919–21.

Rand, D., Rukan, S. B. A., Weiss, P. L., & Katz, N. (2009). Validation of the Virtual MET as an assessment tool for executive functions *Neuropsychological Rehabilitation*, 19, 583–602.

Reeder, C., Smedley, N., Butt, K., Bogner, D., & Wykes, T. (2006). Cognitive predictors of social functioning improvements following cognitive remediation for schizophrenia. *Schizophrenia Bulletin*, 32 (Suppl. 1), S123-31.

Reichenberg, A. & Harvey, P. D. (2007). Neuropsychological impairments in schizophrenia: Integration of performance-based and brain imaging findings. *Psychological Bulletin*, 133, 833–858.

Robbins, T. W. (1990). The case for frontostriatal dysfunction in Schizophrenia. *Schizophrenia Bulletin*, 16, 391–402.

Rodewald, K., Rentrop, M., Holt, D., Roesch-Ely, D., Backenstrass, M., Funke, J., Weisbrod, M., et al. (2010). Targeting Planning and Problem Solving Versus Basic Cognition in Cognitive Remediation for Patients With Schizophrenia. *Schizophrenia Research*, 117, 393–394.

Rodewald, K., Rentrop, M., Holt, D. V., Roesch-Ely, D., Backenstrass, M., Funke, J., Weisbrod, M., & Kaiser, S. (2011). Planning and problem-solving training for patients with schizophrenia: a randomized controlled trial. *BMC Psychiatry*, 11, 73.

Rogers, M. A., Kasai, K., Fukuda, M., & Pletson, J. E. E. (2007). . In J. E. Pletson (Ed.), *Executive deficits in schizophrenia and mood disorder: Similarities and differences.* (pp. 1–40). Hauppauge, NY: Nova Science Publishers.

Royall, D. R., Lauterbach, E. C., Cummings, J. L., Reeve, A., Rummans, T. A., Kaufer, D. I., et al. (2002). Executive control function: A review of its promise and challenges for clinical research: A report from the committee on research of the American Neuropsychiatric Association. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 14, 377–405.

Rushe, T. M., Morris, R. G., Miotto, E. C., Feigenbaum, J. D., Woodruff, P. W. R., & Murray, R. M. (1999). Problem-solving and spatial working memory in patients with schizophrenia and with focal frontal and temporal lobe lesions. *Schizophrenia Research*, 37, 21–33.

Schnirman, G. M., Welsh, M. C., & Retzlaff, P. D. (1998). Development of the Tower of London – Revised. *Assessment*, 5, 355–360.

Scholnick, E. K. & Friedman, S. L. (1993). Planning in context: Developmental and situational considerations *International Journal of Behavioral Development*, 16, 145-167.

Scholnick, E. K. & Friedman, S. L. (1987). The planning construct in the psychological literature. In S. L. Friedman, E. K. Scholnick, & R. R. Cocking (Eds.), *Blueprints for thinking: The role of planning in cognitive development.* (pp. 3–38). New York, NY: Cambridge University Press.

Scholnick, E. K. & Friedman, S. L. (1993). Planning in context: Developmental and situational considerations. *International Journal of Behavioral Development*, 16, 145–167.

Scholnick, E. K., Friedman, S. L., & Wallner-Allen, K. E. (1997). What do they really measure? A comparative analysis of planning tasks. In S. L. Friedman & E. K. Scholnick (Eds.), *The developmental psychology of planning: Why, how, and when do we plan?* (pp. 127–156). Mahwah, NJ: Lawrence Erlbaum.

Semkovska, M., Bédard, M.-A., Godbout, L., Limoge, F., & Stip, E. (2004). Assessment of executive dysfunction during activities of daily living in schizophrenia. *Schizophrenia Research*, 69, 289–300.

Seter, C., Giovannetti, T., Kessler, R. K., & Worth, S. (2011). Everyday action planning in schizophrenia. *Neuropsychological Rehabilitation*, 21, 224–249.

Simon, H. A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, 7, 268–288.

Startup, M., Jackson, M. C., & Bendix, S. (2002). The concurrent validity of the Global Assessment of Functioning (GAF). *British Journal of Clinical Psychology*, 41, 417–422.

Stordal, K. I., Mykletun, A., Asbjørnsen, A., Egeland, J., Landrø, N. I., Roness, A., et al. (2005). General psychopathology is more important for executive functioning than diagnosis. *Acta Psychiatrica Scandinavica*, 111, 22–28.

Stuss, D. T. & Benson, D. F. (1986). *The Frontal Lobes*. New York, NY: Raven Press.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics (5th ed.).* Boston, MA: Allyn & Bacon.

Tucha, O. & Lange, K. W. (2004). *TL-D Turm von London – Deutsche Version [Tower of London – German Version].* Göttingen, Germany: Hogrefe.

Veiel, H. O. (1997). A preliminary profile of neuropsychological deficits associated with major depression. *Journal of clinical and experimental neuropsychology*, 19, 587–603.

Velligan, D. I., Bow-Thomas, C. C., Mahurin, R. K., Miller, A. L., & Halgunseth, L. C. (2000). Do specific neurocognitive deficits predict specific domains of community function in schizophrenia? *Journal of Nervous and Mental Disease*, 188, 518-524.

Viviani, P. (1990). Eye movements in visual search: cognitive, perceptual and motor control aspects. *Reviews of Oculomotor Research*, 4, 353–393.

Wiedl, K. H. & Uhlhorn, S. (2006). *O-AFP Osnabrücker Arbeitsfähigkeiten-Profil* [Osnabrück Work Capabilities Profile]. Göttingen, Germany: Hogrefe.

Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H., & Evans, J. J. (1996). *BADS: Behavioural Assessment of the Dysexecutive* Bury St Edmunds, UK: Thames Valley Test Company.

Wykes, T., Reeder, C., Huddy, V., Taylor, R., Wood, H., Ghirasim, N., et al. (2012). Developing models of how cognitive improvements change functioning: Mediation, moderation and moderated mediation. *Schizophrenia Research*, 138, 88–93.

Wykes, T., Huddy, V., Cellard, C., McGurk, S. R., & Czobor, P. (2011). A meta-analysis of cognitive remediation for schizophrenia: Methodology and effect sizes. *The American Journal of Psychiatry*, 168, 472–485.

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology*, 1, 198–226.

Appendix: Method of the eye tracking and verbal protocol studies

(Description adapted from „A comparison of eye tracking and verbal reporting for tracing planning behavior in a complex scheduling task“, Brüssow, Holt, & Funke, in prep.)

Eye movements were recorded with an EyeLink II system, a head-mounted eye tracker that derives gaze direction from video-based analysis of the pupils. For saccade and fixation detection we used an eye movement velocity threshold of 30°/s, an eye movement acceleration threshold of $8000°/s^2$, and a saccadic motion threshold of 0.1°. All stimuli were displayed at a resolution of 800 x 600 pixels on a 19 inch monitor. Participants were seated 60 cm away from the screen, their head was stabilized with a chin rest during the whole experiment. Viewing was binocular but only the dominant eye was tracked with a sampling rate of 250 Hz. If possible corneal reflection mode was activated to further improve measurement accuracy. A grid of letters was presented before and after each trial block and used to apply a regression correction to the raw data before further analysis in order to reduce systematic deviations in measured gaze direction.

Instructions consisted of a description of the task and a color print of a screen shot taken from the initial training phase of the experiment. Only after participants confirmed that they understood the task the experiment started with the training phase. Each session lasted approximately 60 to 90 minutes. During the experiment response times and response buttons pressed were recorded. Only fixations on the errand descriptions were considered for analysis, which accounted for 67.4% of all fixations (map: 26.5%, schedule: 4.9%, other: 1.2%).

Verbal reporting employed the same stimulus materials as the eye tracking study but presented on a laptop computer with an integrated microphone to record verbalizations. Each session lasted approximately 45 to 60 minutes. Verbal protocol acquisition was less time-consuming because no time was spent on eye tracker calibration. Without the chin rest used in eye tracking the distance to the screen varied between 50 and 60 cm. Following the suggestion of Ericsson and Simon (1993) social interaction was reduced to a minimum by seating the experimenter behind or to the side of the participant. Participants were given the same training phase as in eye tracking to get acquainted with the task. To remind participants to think aloud the instruction given by the experimenter was simply “please keep talking” (cf. Ericsson & Simon, 1993).

Verbal protocols were transcribed and task-relevant information elements corresponding to the interest areas in eye tracking, e.g., location names or task start and end times were coded and time-stamped. This allowed to produce a time series of information usage analogous to the recording of fixations on particular interest areas in eye tracking and for each time interval. The inter-rater agreement for coding the information elements was determined by independently recoding 6 of 25 (24%) reports by a second coder. The proportion of agreement between the coders was 87% and the inter-rater reliability was highly satisfactory with Cohen's kappa of .82. In contrast to eye fixations, the total number of verbalizations notably increased as the task progressed. To eliminate this effect from the analyses presented in this thesis, the counts of verbalizations referring to any given errand were normalized to their relative proportion of the total within each analysis interval.

Declarations in accordance with § 8 (1) b and § 8 (1) c of the regulations for doctoral degrees of the Faculty of Behavioural and Cultural Studies of Heidelberg University.

§ 8 (1) b

I declare that I wrote this thesis independently using only the aids specified and that I have correctly referenced all quotations.

§ 8 (1) c

I declare that I have not submitted this thesis in this or any other form as an examination paper and that I have not submitted it to any other faculty.

Heidelberg, October 2012

Daniel Holt