Abstract

Hidden pathway maze learning tasks (HPMLTs) have been used in neuropsychological research and practice for more than 80 years. These tasks require the use of visual and auditory task feedback signals to learn the order and direction of a pathway, typically within a grid of stepping-stones, or alleys. Hidden pathway maze learning tasks are purported to assess both visuospatial learning and executive processes. The original motivation for the HPMLT paradigm for humans was to reduce a complex tactual planning task to one in which decisions could be directly measured by discrete actions at choice points guided by visual cues. Hidden maze learning paradigms were used extensively throughout the 20th century, initially to study exploratory, anticipatory, and goal-related behavior within the context of memory research, and later as an experimental tool in neuropsychology. Computerization of HPMLTs have allowed for the measurement of different move categories according to the rule structure and ipso facto, clinically meaningful differences in memory and monitoring functions during spatial search and learning. Hidden pathway maze learning tests have been used to understand the cognitive effects of ageing, neurological disorders, and psychopharmacological challenges. We provide a review of historical antecedents relevant to contemporary applications of HPMLTs in neuropsychology. It is suggested that contemporary applications of HPMLTs could be advanced by analysis of component operations necessary for efficient performance that can inform theoretical interpretations of this class of tests in clinically meaningful terms.
Behavior at the Choice Point: Decision making in Hidden Pathway Maze Learning.

“Much of our thinking goes beyond imagery and involves symbolization. Many of our ideas come as insight into complete relationship of things. Maze learning as an ideation process shows that insight may come gradually as the relationships are gradually grasped and coordinated in some symbolic manner by the mind.” (Chase, 1934, pp 438).


Hidden pathway maze learning tests (HPMLTs) require a participant to find the direction of a pathway hidden beneath a small-scale array of alleys, or stepping-stones across repeated trials. These tests differed in structure, requiring the use of a stylus to traverse peg-and-grove alleys, or finger/stylus responses to jump from tile-to-tile within a grid of locations. Hidden pathway maze learning tasks were developed by American researchers in the early decades of the 20th century to investigate spatial reasoning and learning. The stepping-stone variant of the HPMLT paradigm has gained prominence in contemporary neuropsychology as a test of visuospatial learning and executive function (Bowden & Smith, 1994; Walsh, 1985). Hidden pathway maze learning has remained a useful model of cognitive function in neuropsychology because they are relatively brief to administer and are understood easily by healthy children and adults. Furthermore, children with reduced intellectual capacity due to brain damage or disorder have been able to perform HPMLTs because of the simplicity of their rule structure (Schroder, Snyder, Sielski, & Mayes, 2004). Despite the simplicity and brevity of the stepping-stone HPMLT, every move represents a discrete decision point and consequently a large volume of informative data can be obtained within a few minutes of task performance (e.g. the Groton Maze Learning Task (GMLT):
Snyder, Bednar, Cromer, & Maruff, 2005). Computerized versions of HPMLTs have allowed for greater analytical scope than manual predecessors of the task because every move can be recorded, classified, and replayed. When pre-training and alternate pathways of the task are available, performance by healthy, or brain damaged participants shows little evidence of practice effects, which is an extremely rare quality in neuropsychological assessment. These factors make HPMLTs ideal for use in repeated measures research designs (e.g. Snyder, Jackson, Piskulic, Olver, Norman, & Maruff, 2008a). These characteristics have allowed researchers to build brain-behavior models on the basis of general performance effectiveness (i.e. from analysis of total scores) and also by allowing analysis of the component cognitive processes as well as their integration (Pietrzak, Maruff, & Snyder, 2009a; Thomas, Reeve, Pietrzack, & Maruff, 2013). The HPMLT model has been a useful tool in neuropsychology because impairment and improvement in spatial reasoning and memory functions can be reliably measured over the course of normal development, and across a range of neuropsychological disorders and treatment interventions (Pietrzak, Cohen, & Snyder, 2007; Snyder et al., 2008a; Thomas, Reeve, Fredrickson, & Maruff, 2011).

While computerized HPMLTs are relatively new to neuropsychology, maze learning has been used to study adaptive goal-related behavior since the formalization of psychological science (Hamilton, 1911; Melrose, 1922; Ruckmick, 1921; Small, 1901). Current concepts of executive function (Lezak, 1982) and spatial memory (Kessels, deHaan, Kappelle, & Postma, 2001; O’Keefe & Nadel, 1978) were derived from maze learning methodologies designed to investigate behavior that was structured toward a physically and temporally remote goal (Melrose, 1922; Ruckmick, 1921). The underlying argument of this review is
that historical investigations of motivation and learning through the use of HPMLTs are continuous with contemporary investigations of spatial memory, and executive functions such as error monitoring. Analysis of the development of maze learning methodologies is therefore important for understanding the theoretical relevance of HPMLT paradigms for neuropsychology.

In the first section of this review examines the historical considerations through which the stepping-stone variant of HPMLT was developed and analyzed. It is asserted that HPMLTs were developed to investigate the processes of cognitive integration and co-ordination in goal-directed learning and these considerations were the precursors to contemporary accounts of executive function and spatial navigation. The second half of the review considers the modern uses of HPMLTs, specifically focusing on use of versions of the stepping-stone maze in neuropsychology. It is argued that qualitative and quantitative analysis of performance is necessary for characterizing change, or difference in cognitive abilities in clinically meaningful terms.

1.2. Hidden pathway maze learning: Historical considerations.

In the introduction to the first edition of the *Journal of Experimental Psychology*, Trumball Ladd (1894) outlined the historical background to the objectives of the American Psychological Association (APA) by asking the question: “How shall we regard the science of mental life as related to the methods and conclusions of the most nearly allied physical sciences, to philosophy, and to human action and character?” (Trumball Ladd, p. 3). One of the aims of the APA at that time was to establish scientifically credible protocols and methods for investigating complex mental and behavioural processes in non-teleological and
non-metaphysical terms (Hatfield, 2002; Trumball Ladd, 1894). Maze learning paradigms were adopted early in psychological science because they specifically addressed the problem of how an animal determines the nature of an environment for ecologically salient rewards (such as food) that are not immediately present, and therefore are not contiguous determinants of behavior (Small, 1901). The HPMLT paradigm met the standards required of an objective factual science from which intelligent, problem-solving behaviour could be interpreted with regard to a cross-species “natural history of the mind” (Small, 1901, p. 222).

From the beginning of comparative psychology (e.g. Darwin, 1871), the cognitive functions of animals and humans, especially children, in maze learning were considered sufficiently similar to regard common behaviours as reflecting a shared intellectual foundation (Hatfield, 2002; Lockard, 1971; Tolman, 1938; Small, 1901). Specifically, researchers in comparative psychology based their models on the assumption that animals and humans rely on common trial-and-error methods for problem-solving (Small, 1901, Hamilton, 1911). The trial-and-error exploration required by maze learning was therefore well suited to the objectives of early psychological science for investigating selective processes underlying adaptive behavior on the proposition that discrimination learning was the “basis of each and every act” (James, 1890 cited by Ruckmick, 1921, p. 3-5; Stephens, 1935). Tests of discrimination learning were also central to advancing theories that defined intelligence according to the ability to monitor performance on a complex task and adapt responses for the efficient pursuit of an end goal (Binet & Simon, 1909, cited in Peterson, 1922, p. 370; Porteus, 1918). Decisions made at each choice-point in a maze provided units of behaviour that could be analysed according to the ability to learn correct turns in the path from a blind alley (Brown
The maze context was important because it allowed for the study of behaviour when there is only partial knowledge of the environment. It was not simply a method for studying the learning process, but a method for measuring how an animal determines the nature of the environment (Tolman, 1948). Researchers investigated how expectancies of a task solution developed by the extent to which ineffective ‘habits’ of entry into blind alleys were eliminated (Brownwell, 1939; Lashley, 1923; Muenzinger, 1927; Pepper, 1934; Perry, 1918; Peterson, 1922; Tolman, Hall, & Bretnall, 1932). Thus, in animal studies maze learning was considered an appropriate operational definition for higher cognitive processes because it involved “the capacity to take in a large situation, and by response to remote conditions to delay action and inhibit motor tendencies aroused by immediate stimuli until they have been coordinated with the larger circumstance” (Peterson, 1922, p. 387). Maze learning was therefore regarded as a measure of intelligent learning in which the experience of previously encountered elements of the maze was used to guide decisions toward a future state of reward (Peterson, 1922), such as food in animals, or task completion in humans.

In general, maze learning was described as a complex process of eliminating and integrating behavioral tendencies, or representational elements of a cognitive map, in learning a maze pattern (Hull, 1932, 1934a, 1934b, 1938; Tolman, 1938, 1948, 1959).

One important characteristic of the rodent maze learning paradigm was that it could be applied largely unchanged to study problem solving in humans (Hamilton, 1911; Miles, 1927; Warden, 1924a, 1924b; 1925). Early studies of maze learning in humans were conducted using small-scale table-top mazes with
subjects blindfolded and using their finger to process tactual cues (Chase, 1934; Koch & Ufkess, 1926; Melrose, 1922; Scott, 1930; Warden, 1924b). Blindfolding in humans was argued to allow for direct comparison of the cognitive processes involved in maze learning between animals and humans: deprivation of vision made maze learning dependent on motor cues, as seemed to be the case for rodents (Melrose, 1922; Warden, 1924b). According to Small (1901), tactual-motor experience was fundamental to spatial perception. It was generally argued that by blindfolding, the primitive basis of experience in kinesthetic and motor cues could be studied directly without the overlay of latter evolving visual and linguistic processes (Brownwell, 1939; Peters & McClean, 1935; Warden, 1924b).

1.3 The development of the hidden pathway maze learning paradigm

Although the use of blindfolding was intended to make human and animal maze learning paradigms similar, the deprivation of vision and forced reliance on tactual-motor information was not a normal context for assessing learning and cognition in humans. For example, the participants in Jensen’s (1934) study required approximately 22 trials and participants made over 200 errors, with or without punishment by electric shock, to negotiate a blindfolded 30 choice point maze (e.g. see Figure 1.3.1).

The blindfolding method used in maze learning studies was observed to cause confusion in humans with regard to orienting to the maze for discovery of the pathway direction (Gould & Perin, 1916; Peterson, 1920). Errors arising from the
disorientation in blindfolded maze learning were interpreted as ‘frustration’ leading to ‘non-rational’ behaviour such as repeating mistakes and perseverating at points of disorientation (Peterson, 1920; Barker 1931). Consequently, apparatus emerged that were designed to reduce the artificiality of tactual maze learning tasks for humans and with restricted visual information without blindfolding, but which retained the focal construct of maze learning as a method for studying ‘rational behaviour’. For example mazes were developed that allowed indirect visual control such as a map, or ‘two-story duplicate’ maze (Miles, 1927- see Figure 1.3.2), or the use of hidden stops for hiding the path in a peg-groove stylus maze (Carr, 1921- see Figures 1.3.3 to 1.3.4).

In the Miles (1927) maze, the maze layout could be seen from the top. However, the top view maze was useful only for negotiating the bottom maze hidden from view if some form of visuo-tactile transformation could be used to bring the two mazes into alignment. In the Carr (1921) maze, the entire maze problem is revealed, but slotted grooves hold the stylus in place (Figure 1.3.3), and physical stops beneath the visible apparatus signal entry into blind alleys (Figure 1.3.4). In both examples, the maze layout could be seen, but required transformations such as image rotation (Miles, 1927), or memory for the location of blind alleys (Carr, 1921). The innovation of partial visual cues to the maze layout overcame the impedance of blindfolding to maze learning while preserving the complexity of the task as a spatial learning paradigm. In studies from the early decades of the 20th century, the role of visual guidance on acts of skill in a motor task was discussed in theoretical and practical terms (see also, Koch, & Ufkees, 1926; Warden, 1924a). Mainly, researchers were concerned with how symbolic knowledge of space, or ‘ideation’ is built from the interaction of sensory cues with
intellectual processes (Carr, 1921; Chase, 1934; Cox, 1928; Perrin, 1914; Sartian, 1940; Tilborg, 1936; Warden, 1924a, 1924b, 1925). Chase (1934) concluded that in ideational learning all senses are utilized in order to coordinate the central organization of ideas beyond language and imagery.

The use of indirect visual guidance, as a practical concern in maze learning paradigms, presented methodological problems for designing apparatus for use in humans that had not been so important in blindfolded mazes. For example, in a blindfolded maze, the number and length of turns at a choice point are not apparent from the outset of the task, as they are when the maze layout can be scanned visually. Only tactual and kinaesthetic exploration of the alleys one-at a time can reveal the spatial array, upon which future choice point selection can be made when a location is revisited in future trials (Cox, 1928). However, in mazes that are partially revealed, all alleys and turns are immediately present for visual selection at any given choice point. It was recognized that the length and number of turns, as well as the direction to the visible end-goal could influence the probability of selecting an alley. As focus turned from the number of errors or trials to task completion in maze learning to the probability of selecting particular turns (Hill, 1939; Hull, 1932; 1934a, 1934b; Tolman, 1938), consideration of the maze layout became more important (e.g. Jones & Yoshioka, 1938; Warden,
Researchers consequently designed standardized maze apparatus in which every alley was of equal length and with the same number of blind alleys with a single correct alley available for selection at every choice-point (Griffith, 1931; Warden, 1924a). By standardizing the configural aspects of maze pathways, such as the number and direction of turns, and the length of the alleys, the task could be analysed as homogenous units with equal potential for selection (Warden, 1924a: see figure 1.3.5).

FIGURE 1.3.5 ABOUT HERE

The introduction of electrically mediated visual and auditory feedback signals allowed development of maze learning paradigms in which blind alleys could be signalled symbolically, for example by activating a light or a buzzer when a blind alley was entered, rather than relying on physical stops to indicated dead ends (Barker, 1931; Faber & Berman, 1938; Gilbert, 1934; Gurnee, 1938; Razran, 1936, 1939). Furthermore, the ability to signal errors at every move made the need for physical alleys redundant. The maze could be represented by a field of locations, each of which could be selected one-at-a-time like stepping-stones, with immediate feedback for each step as correct or incorrect. Hence, movement along a series of stepping-stones could be equated with the forward movement or turns allowed by the alleys in a tactual maze. Contiguous points in the surrounding grid defined movement options spatially. Movements in the maze were therefore standardized because every step in the maze pathway gave rise to equal and identical response options for horizontal and vertical moves in a two-dimensional space. In electrical mazes, each location was connected to a contact plate, or screw head contacts.
wired into an electrical circuit. The circuit was designed so that true points in the pathway gave different feedback to correct and incorrect locations with lights, buzzers, clickers, or shocks, but the pathway itself was not visually revealed (see Figure 1.3.6).

FIGURE 1.3.6 ABOUT HERE

Barker, (1931: Figure 1.3.6), Griffith (1931), Razran, (1936, 1939), Faber & Berman, (1938) and Mann & Jewell (1941) describe similar electrical stepping-stone apparatus that differed in grid size, pathway length, and the form of error feedback. All these versions relied on step-by-step trial-and-error exploration to discover a pathway within a grid of stepping-stone tiles. Screw points representing each location were wired into a circuit that activated correct and incorrect feedback on contact with a thimble stylus. The pathway could be changed by rewiring the order of screw points in the apparatus, although these tasks were not used experimentally, except as a distracter task (Razran, 1936, 1939) until Mann & Jewell (1941).

The advantage of the stylus grid maze over traditional alleyway maze paradigms was that by presenting the maze as a homogenous grid of stepping-stones, the pathway to be solved could not be mapped visually onto alleys. By restricting the field of visual attention to the immediate location in the pathway, the maze could only be discovered through individual trial and error movements at each choice point, rather than by forward visual exploration of alleys. In this way, a complex task was reduced to a series of simple decisions at standardized choice-points, revealed in manual choices based on the ability to develop and apply a mental representation of the task. Electrifying the feedback also removed the
influence of the experimenter on performance and allowed for automatic recording of each move (see Ruckmick, 1927, for an electrified version of the Warden maze).

Mann & Jewell (1941: Figure 1.3.7) first described a rule structure to guide performance in the stepping-stone maze. A rule structure is useful in stepping-stone maze designs because moves are not physically restricted to contiguous alleys as they are in the alley stylus mazes: In stepping-stone mazes moves can be made anywhere in the grid. Therefore to measure maze choice-point behavior consistent with a stylus maze, rules must be imposed to limit possible choices to immediate contiguous locations.

FIGURE 1.3.7 ABOUT HERE

The development of the stepping stone maze was a technological advance in the automation and standardization of apparatus in psychology. Most studies describe the stepping-stone maze in practical terms of attempts to control extraneous aspects of the task, such as biased expectation of turns, or frustration from disorientation in other maze formats (Barker, 1931: Faber & Berman, 1938). The ability to record and classify every move in the task conferred a statistical advantage for analyzing the units of complex behavior. The elements of cognition could be studied as individual decisions in choice-point behavior, and as a unified whole, such as goals, expectations and according to factors that influence immediate behavioral determinants (Jones & Yoshioka, 1938). Theories of complex volitional learning behavior developed throughout the 20th century did not focus on the stepping stone maze (e.g. Tolman, 1939) because it was rarely used until the latter half of the 20th century. However the potential for use of HPMLT paradigm in theoretical inquiries of spatial reasoning is
apparent from quantitative and qualitative analysis of maze performance in at least one study (Jones & Yoshioka, 1938).


Animal and human performance in hidden maze pathway learning was usually expressed as a reduction in errors over trials also termed a learning curve (Gould & Perin, 1916; Lambert & Ewart, 1932; Thurstone, 1933). However, researchers of both animal and human behavior were also interested in qualitative parameters of the learning processes, such as the rate of elimination of specific choice points in relation to the end goal (Hull, 1934a, 1934b; Jones & Yoshioka, 1938; Peterson, 1920). It was observed that some turns were harder to learn than others, therefore in accounting for maze behavior it was important to understand how habits and expectancies of a maze solution changed with experience (Brown & Buel, 1940; Hill, 1939; Hull, 1934, Jones & Yoshioka, 1938; Meunzinger, 1927; Meunzinger & Vine, 1941; Tolman, 1932, 1938; Voeks, 1948, but see Mann & Jewel, 1941). Rather than simply analyzing learning curves, it was common to analyze qualitatively distinct categories of errors that related to different types of behaviors, such as anticipation of turns, or avoiding correct turns that lay in the opposite direction to the end goal (Brown & Buel, 1940; Jones & Yoshioka, 1938). By analyzing how different errors changed with experience it was possible to describe how ‘schemes of control’ are built through experience that distinguish learning efficiency (Gould & Perin, 1916, pp 142) and to develop theories of complex behavior (Hull, 1934; Tolman, 1938). Theories of maze learning were concerned with anticipatory and avoidance behavior that could not be simply explained by association learning, but required considerations of goals or purpose.
through qualitative and quantitative analysis (Hull, 1934a; Pepper, 1934; Perrin, 1923; Tolman, 1932).

Although maze learning paradigms were used extensively in the first half of the 20\textsuperscript{th} century, only a few studies used the stepping-stone maze before the 1960’s within the theoretical framework of expectancy and change in motivated behavior (e.g., Jones & Yoshioka, 1941; Mann & Jewell, 1938: but see Gurnee, 1938; and Razran, 1936, 1939 for quantitative studies of reinforcement learning in the stepping stone maze).

In one study, a comprehensive analysis of maze learning behavior in adolescent humans was based on rodent models of maze learning (Jones and Yoshioka, 1938). Jones and Yoshioka studied 159 adolescents from 11 –to -15 years on a 5 x 5 stylus maze with 22 and 23 steps (see figures 1.4.1 -1.4.3) with (1) no feedback, (2) a buzzer, or (3) a light on the fourth and subsequent trials signaling errors.

\textbf{FIGURE 1.4.1 ABOUT HERE}

\textbf{FIGURE 1.4.2 ABOUT HERE}

\textbf{FIGURE 1.4.3 ABOUT HERE}

Results of this study revealed that the nature of task feedback had no effect on the reduction in errors. In all conditions, the children achieved error free performance after 13 to 15 trials. Interestingly, the authors noted that many of the errors made by children were incorrect responses made in the general direction of the end point, where the path continued laterally or turned away from the goal. Other frequent errors were with regard to anticipating turns, and to avoid the periphery of the maze. Individual differences were shown in the tendency to
retrace and repeat errors and tended to occur most frequently when the correct pathway turned away from the goal. Jones and Yoshioka (1938) described these errors as reflecting (a) goal impetus; (b) drift impetus in the direction of the section of the current pathway; and (c) local forward momentum where continuous trajectory is made until an error, where (a) and (b) give rise to anticipatory errors. This study is important because it was the first to characterize comprehensively the nature, frequency, and change in errors in hidden pathway electrical stylus maze learning (see Brown & Buel, 1940, for a discussion of the findings of Jones & Yoshioka, 1938).

Errors such as retracing or repeating the same error were often reported in early animal and human maze learning studies (e.g. Peterson, 1920; Hill, 1939). These errors were described as ‘perseveration’ and were considered to define confused and irrational behavior (Gould & Perin, 1916; Peterson, 1920, 1922). Precursors to modern neuropsychological interpretations of maladaptive behavior were evident from the time of Binet and Simon (1909), Hull (1934a, 1943b), Porteus (1958) and Tolman (1959), in so much as they considered variability in adaptation in the maze learning context. However, analyses of repeating errors in the stepping-stone maze were not conducted until decades later when it was considered an important measure of perseveration behavior in neuropsychological studies (Karnath, Wallesch, & Zimmerman, 1991; Milner, 1965).

2.1. Hidden pathway maze learning and modern neuropsychology.

Maze learning tasks were considered useful for studies of cognition in patients with different focal brain lesions because of the extensive use of this paradigm in animal learning studies (Orbach, 1955, 1959). Furthermore the processes involved in different spatial tasks were well established compared to other areas of
neuropsychological research (Semmes, Weinstein, Ghent & Teuber, 1955). Non-verbal tasks, such as maze learning, were useful in neuropsychological experiments because they could be used to study patients with impairments in aspects of language that were a common presentation after focal cortical injury (Teuber, 1966). However, the relationship between brain function and aspects of spatial processing, such as taxon navigation, object-location learning, or topographical memory was not well understood (Kessels et al., 2001; Olton, 1985), and consequently the neuropsychological interpretations of impairments in hidden pathway maze learning have been speculative.

2.2. Maze learning and lesions of the temporal lobes.

Hidden pathway maze learning paradigms were most commonly used in neuropsychological research after the studies of Milner, (1965), Teuber (1966), and the Oxford group (Newcombe & Russell, 1969) investigating the role of the hippocampus in memory processes. The importance of hippocampal and medial temporal lobe brain structures to memory encoding and retrieval was exemplified in patient H.M., who displayed severe verbal and spatial anterograde amnesia following bilateral mesial temporal lobectomy (Milner, Corkin, & Tueber, 1968). Despite this dense amnesia, H.M.’s short term-spatial memory capacity, as measured by Corsi Block span, was within the normal range (Milner, 1971). H.M. and other patients with focal damage to the right medial temporal lobe appeared to have difficulties learning new information that was beyond the span of immediate memory (Milner, 1965, 1971; Milner et al., 1968). To investigate the extent to which focal hippocampal lesions disrupted learning in the spatial domain, Milner needed a task with a temporal, or sequential ordering component that was beyond the span of immediate spatial memory and therefore required cumulative learning
Milner used a HPMLT that was similar in design to the Mann & Jewel (1941) maze. The Milner maze and later-developed Oxford HPMLT were simpler than the Porteus or Elithorn perceptual mazes because they did not require any development or evaluation of potential global solutions for problem-solving. They were also simpler than human scale mazes that were navigable by the use of maps for orienting in a landscape (Milner, 1965; Semmes et al., 1955). The Semmes et al., (1955) maze required the translation of a 2-dimensional paper map to the location and orientation within a physical landscape. Furthermore, the rule structures of HPMLTs were simple and could be understood easily even by patients with specific disturbances of language (e.g. arising from focal lesions of Broca’s area). For example, Milner (1965) showed that patients with Broca’s aphasia could perform the HPMLT within normal limits (Milner, 1965).

In Milner’s HPMLT, individuals were required to locate a 29-step pathway with 11 turns hidden within a 10 x 10 matrix of stepping stones, or locations to complete a single learning trial. The individual had to find the path traversing contra-lateral corners of the grid using move rules for the selection of each contiguous location using a stylus pen (see Figure 2.2.1). The rules of the Milner HPMLT included: go back to the preceding bolthead location when an error is signaled (by an auditory ‘click’); do not retrace sections of the correct pathway (or do not backtrack on the pathway); and do not move diagonally (Milner, 1965).

**FIGURE 2.2.1 ABOUT HERE**

Milner’s criterion for task completion was the individual obtaining three successive error free trials, or to complete a total of 25 trials in two blocks given daily, or until criterion was reached. On the Milner HPMLT, the patient H.M. was
unable to learn the maze within any single session and also unable to reduce the number of errors made learning the same pathway even after 215 trials across days. The inability to learn the pathway on the Milner maze was unrelated to any perceptual or intellectual disability arising from his brain injury (Milner et al., 1968). Subsequent studies interpreted the deficits in performance on HPMLTs observed in patients with right retro-Rolandic cerebral damage as topographical amnesia that was independent of spatial perception or general intelligence (De Renzi, Faglioni, & Villa, 1977; Newcombe, Ratcliff, & Damasio, 1987; Newcombe & Russell, 1969; Ratcliffe & Newcombe, 1973).

The use of HPMLTs to study the effect of lesions or impairment to the hippocampus in humans was well established by the mid 1970’s, mostly using The Oxford Stylus Maze, or the Stylus Maze Test (Newcombe et al., 1987; Newcombe & Russell, 1969; Ratcliff & Newcombe, 1973), or HPMLTs that used similar designs (Talland, Hagen, & James, 1967; Talland & McGuire, 1967). The Oxford HPMLT was similar in design to peg and groove mazes like the Jones & Joshioka (1938) and the Sartian (1940) HPMLTs. The Oxford Stylus HPMLT consisted of an 8 x 8 grid of raised blocks where individuals traced a path in an alley between locations from a start point to a finish point with auditory and visual feedback signals given when the path was traced onto an incorrect location. When errors were signaled the individual returned to the last choice point before moving again. Task A was a 6 choice point path in which the examiner demonstrated the path and the participant was required to relocate the pathway to a criterion of 3 error free trials, or 26 trials. This was used to allow training and practice trial to ensure task understanding (Newcombe & Russell, 1969). Task B had 10 choice points and the pathway was not demonstrated, therefore it had to be discovered by trial and error (see Figures 2.2.2(1) and (2)).
The Oxford HPMLT was developed as part of a project to study the long-term cognitive effects of focal brain injuries, including those affecting the medial temporal lobes and posterior parietal cortex (Newcombe & Russell, 1969). The structural similarities between the Milner and Oxford HPMLT are unlikely to have been accidental, given the reported communication between the Montreal and Oxford groups regarding common research interests (Newcombe & Russell, 1969). The Oxford HPMLT contained 10 steps with 6 turns. Milner et al., (1968) reported that H.M. showed some evidence of slow memory acquisition and retention on a modified version of the Milner HPMLT that included 8 steps with 3 turns, but was unable to learn a HPMLT with 10 steps and 4 turns (Milner et al., 1968). H.M.’s ability to learn the smaller HPMLT was likely due to the capacity to encode the pathway within working memory, which was intact. Although it is common to examine the effect of increasing task difficulty on aspects of performance in neuropsychological tests (e.g. De Luca et al., 2003; Luciana & Nelson, 1998, 2002; Welsh, Pennington & Grossier, 1991), only two studies to date have examined the effect of increasing task difficulty, such as grid size and pathway length on a stepping-stone HPMLT (Mathewson, Dywan, Snyder, Tays, & Segalowitz, 2008; Thomas et al., 2011). Both studies found that manipulation of HPMLT pathway length/grid size was associated with qualitative changes in performance in children (Thomas et al., 2011), as well as neurological differences between older and younger adults (Mathewson et al., 2008). The results of these studies are discussed later in this review (see Section 3.3.)
Using the Oxford HPMLT, Newcombe and Russell (1969) found evidence of
dissociation between visual perception that was impaired in patients with occipital
lesions, and maze learning that was specifically impaired in groups with temporo-
parietal lesions, but not in patients with occipital lesions. In accordance with Milner
(1965), Newcombe and Russell (1969) argued that performance of the different lesion
groups on the HPMLT could be dissociated such that impairment due to right
hemisphere lesions could be interpreted as a spatial cartographic problem in posterior
parietal groups, and a mnestic disturbance in patients with right temporal hippocampal
excision.

2.3. Maze learning and lesions of the frontal lobes.

The importance of the frontal lobes to cognitive models of problem solving
became evident very early in the history of neuropsychology (Bigelow, 1850). For
example, Teuber (1963, 1966) speculated that lesions of the frontal lobes disrupted
the coordination of cognitive processes that depended on other brain structures.
Experimental paradigms used to describe the effects of lesions involving the frontal
lobes exploited dissociations between the ability to evaluate errors in performance and
to utilize information from errors in planning future actions (Konow & Pribram,
1970). Another way of considering this was as the ability to structure behavioral
sequences and monitor their outcomes (Luria, 1963, 1973). Impairment in this type of
complex adaptive behavior was observed commonly in patients with frontal lobe
lesions.

In order to study disturbances in adaptive behavior in patients with lesions
involving the frontal lobes, maze tasks were utilized often because optimal
performance on HPMLTs required multi-step forward planning. Paper-and-pencil
perceptual mazes that depended on visual analysis were used to study non-verbal executive functions, as ‘planning and monitoring’ tasks (e.g. Elithorn maze: Benton, Elithorn, Fogel, & Kerr, 1963; Elithorn, 1955; Porteus maze: Landis & Erlick, 1950; Porteus 1965). In general, the Porteus and Elithorn mazes measured the ability to plan the most efficient route through a visual array of alleys or dots that were connected through a series of grids that could be increased in size in order to increase the complexity of decision-making demands. Route finding efficiency in the Porteus maze was measured by the number of entries into blind alleys, and by the number of dots traversed in the Elithorn maze (Benton et al., 1963; Landis & Erlick, 1950). Failure to follow the task rules was also an important measure of the ability to maintain task set, or rules during maze performance (Benton et al., 1963; Landis & Erlick, 1950). Both tasks measured rule break errors, such as deviating from the pathway boundary and backtracking (Benton et al., 1963; Landis & Erlick, 1950). Neuropsychological studies indicated that performance on paper-and-pencil perceptual maze tasks was impaired in people with focal lesions involving the right/bilateral frontal lobe, as well as the right parietal lobe, suggesting that both planning/monitoring and spatial processing functions were necessary for optimal performance on these tasks (Elithorn maze: Benton et al., 1963; Elithorn, 1955; Porteus maze: Landis & Erlick, 1950; Porteus & Kepner, 1944). Interestingly, though, the theoretical models of maze learning, as a measure of adaptive behavior and motivation that had been developed in earlier decades (e.g. Tolman, 1959; Hull, 1934a) were mostly disregarded in neuropsychological studies. Instead, brain-based models of maze learning focused on the construct that a task was designed to measure, such as learning or planning (e.g. Lashley, 1943; Teuber, 1966).
While performance on perceptual mazes was disrupted by focal lesions of the frontal and parietal lobes, qualitative analyses of the errors made by these different patient groups were not conducted and so it is not possible to determine post-hoc whether the nature of performance impairment differed across clinical groups (Benton et al., 1963). This is perhaps because the errors made on perceptual mazes could not be understood easily in terms of prevailing neuropsychological models of that time (e.g. Landis & Erlick, 1950). For example, solving perceptual mazes required a co-ordinated series of steps that take into consideration the entire problem space, as well as evaluation of possible global solutions. Individual moves were inextricably tied to a multi-step plan of action. Moves that did not conform to the ideal solution could not be easily understood according to a theoretical account of the discrete behavioral processes of maze learning. As HPMLTs were designed to reduce each move to a local series of options within the problem space, each move could be analyzed with regard to behavior at every choice point (Jones & Yoshioka, 1938). In this way, maze learning could be regarded as a series of conceptually similar choice-point sub-problems. Each move constituted a discrete decision unit in a local problem for finding the next step, and which was guided by the overall goal of learning the pathway (Griffith, 1931). The HPMLTs retained the focal construct of multi-step forward planning, and spatial learning, but allowed for analysis of each move in each sub-problem in theoretical terms. In contrast to perceptual mazes, HPMLTs were amenable to qualitative and quantitative error analysis, which could be used to guide clinical interpretations of focal brain injuries (Milner, 1965). The HPMLTs were also useful because patient groups with different cortical lesions were often impaired in both memory and executive functions but it was difficult to identify the
contribution of these different impairments to poor performance on a single neuropsychological task. Milner (1965) was the first to argue that the effects of different cortical injuries on problem solving and learning could be dissociated using analysis of error types on the stepping-stone version of HPMLT.

Milner recognized that her HPMLT was sensitive to injuries of the frontal lobe but that this sensitivity depended on analysis of the nature of errors made by patients during pathway learning. For example patients with focal injuries involving the frontal lobe made more errors than controls or patients with aphasia (arising from lesions of Broca’s area), in that they revisited locations signaled as incorrect more often, back-tracked along the pathway or jumped to non-contiguous locations (Milner, 1965). Despite the higher rate of these rule break errors, patients with brain injuries involving the frontal lobes could report that they knew their errors were in violation of the task rules. Subsequent studies have similarly found that failing to return, and within search (often termed perseverative errors) are more common than other rule break error types in HPMLTs (Karnath et al., 1991; Matson, Berk and Lucas, 1997; Milner, 1965; Pietrzak et al., 2007; Snyder et al., 2005a; Snyder Maruff, Pietrzak, Cromer, and Snyder, 2008b; Steinberg, Chaikelson, and Schwatzman, 1983; Thomas, Snyder, Pietrzak, Jackson, Bednar and, Maruff, 2008). This pattern of errors has also been observed in children, and occurs most often when the pathway turns in a direction that is against the end goal, or in an unexpected direction (Thomas, personal observation). Rule breaking behavior could be discussed from several theoretical perspectives of error monitoring, behavioral regulation, or ‘fixated’ behavior. However theoretical interpretations of rule breaking behavior and its causes has not been widely discussed in relation to HPMLTs.
While the Milner HPMLT had been developed for study of cognitive impairment in patients with lesions involving the hippocampus, it was applied more broadly to understand cognitive impairment in patients with neuropsychiatric conditions associated with deficits in executive functions, such as schizophrenia, obsessive compulsive disorder, as well as in patients with frontal lobe lesions (Behar, Rappoport, Berg, Denkla, Mann, Cox et al., 1984; Canavan, 1983 [Oxford HPMLT]; Head, Bolton, & Hymas, 1989; Karnath et al., 1991 [modified HPMLT]; Matson et al., 1997 [Austin HPMLT]; Rettew, Cheslow, Rappoport, Leonard, & Lenine, 1991; Steinberg et al., 1983; Wallesh et al., 1990 [modified HPMLT]). In most studies using the Milner HPMLT and its variants, qualitative and quantitative analysis of errors were used to guide interpretation of impaired performance in the clinical group under investigation. For example, following Milner (1965) it was common in the studies cited above to compare the rate of specific errors in HPMLT made by different patient groups, such as repeating errors, breaking rules and to differentiate these from errors that comply with the rules, but probably reflect difficulty learning the pathway over trials. Typically errors were classified as qualitatively different forms of behavior that reflected rule-following, or rule-breaking and were interpreted as distinct disturbances in brain areas associated with error evaluation, and spatial learning and memory. Unfortunately, most studies using the Milner HPMLT that have reported separate measures of rule break and spatial memory errors have not included performance from any of the control groups they also studied (for exceptions see: Head et al., 1989; Hymas, Lees, Bolton, Epps, & Head, 1991; Rettew et al., 1991). Consequently, no guidelines had been developed to classify normal from impaired performance for the different types of errors on the HPMLT (Bowden & Smith, 1994). Subsequently, researchers using HPMLT have shown the sensitivity to
detect change or difference in cognition according to analyses of error types reflecting the spatial learning and rule breaking aspects of performance in terms of effect sizes (Pietrzak et al., 2009a; Pietrzak et al., 2009b; Thomas et al., 2008).

2.4. The Austin Maze

The Austin HPMLT (see Figure 2.4) is a version of the HPMLT Milner maze designed for use in clinical neuropsychological settings as a test of cognitive functions that depended on the frontal lobes (Darby & Walsh, 2005; Walsh, 1985). The popularity of the test in neuropsychology was because it had proved to be a sensitive test of cognitive dysfunction, at a time when there were few tests that were specific to frontal lobe injuries (Burgess et al., 2006).

Darby and Walsh (2005) describe several case studies that highlight the sensitivity of the Austin HPMLT to injuries involving the frontal lobe. In these cases patients with lesions or injury involving the frontal lobes typically made persistent errors on the HPMLT showing that like H.M., they were unable to learn the maze pathway. However H.M. and other patients with unilateral or bilateral hippocampal, temporal, or parietal injuries made almost no repeated errors that broke the task rules. H.M.’s inability to learn the pathway was not due to a failure to follow the task rules, but was attributed to a disturbance of sequential behavior that prevented the acquisition of the pathway in memory (Milner, 1965). Unlike H.M., the inability to learn the pathway in patients with frontal lesions was due to their inability to adjust their behavior on the basis of the errors made, whether due to disinterest in avoiding
errors, or some other dysfunction (Milner, 1965). This inability was considered by Walsh to indicate some disability to regulate purposive behavior (Darby & Walsh, 2005, pp 141, 149). The limited capacity to change behavior on the basis of errors, or repeating the same error continually despite negative feedback, was observed in patients who had suffered traumatic head injury and in adults with chronic alcoholism (Darby & Walsh, 2005). Walsh was careful to emphasize that these deficits reflected a problem of error utilization, rather than error evaluation, because all of their patients demonstrated that they knew their errors were occurring but despite this did not modify their behavior in accord with this knowledge (Darby & Walsh, 2005, pp 168). Hence, this impairment was framed as a difficulty in implementing behavioral plans, not impairment in awareness, or in the assessment of the significance of the errors that broke the task rules. On the HPMLT this failure to learn from errors was observed as a rapid reduction in errors across the initial learning trials with the persistence of a small number of errors across the remainder of trials with performance never becoming error-free (i.e. defined by Walsh as 3 errorless trials) despite individuals being allowed to continue learning the maze indefinitely. It appears that this criterion of error-free performance within 20-30 trials (Walsh, 1985) was selected because Walsh (1960) also reported that patients who had undergone frontal leucotomy also made rule break errors similar to those that had been observed by Milner (1965). However, Walsh found that the rate of rule break errors in patients with frontal lobe leucotomy was not significantly greater than controls (Darby & Walsh, 2005 pp, 149, 167). Studies have consistently shown that the rate of rule break errors is typically much lower than the rate of pathway memory errors, except in patients with severe frontal lobe injuries (Behar et al., 1984; Canavan, 1983; Milner, 1965; Rettew et al., 1991; Steinberg et al., 1983). Inter-individual differences in HPMLT errors can be
shown using metrics that consider the relative rate and variability of differences within and between groups (i.e. effect sizes), than the raw rate of different errors (Snyder, et al, 2005a; Snyder et al., 2005b).

Bowden and Smith (1994) criticized the use of a criterion requiring stable error-free performance to define difficulties of error utilization in HPMLTs, from both a statistical and theoretical perspective. First, stable error-free performance could be achieved in patients with injuries involving the prefrontal cortex, albeit after many more trials than required by healthy controls (Milner, 1965). Second, in some healthy adults error-free performance required as many as 50 trials (Bowden et al., 1992), and therefore failure to reach criterion in 20 or 30 trials (Walsh, 1985) could not be considered a reliable measure of cognitive impairment. Finally, Bowden and Smith (1994) showed that the rate of learning (i.e. reduction in errors) over the initial 10 trials on the Austin HPMLT was highly predictive of the number of errors on subsequent trials made by both healthy adults and patients with heterogeneous brain lesions (Bowden & Smith, 1994). Although it was traditional to examine maze performance in terms of trials to a criterion of error free completion, after the studies of Bowden, it became increasingly common to examine errors over a pre-specified number of trials (e.g. Barker, Greenwood, Jackson, & Crowe, 2005; Morrison & Gates, 1988; Österberg, Karlson, & Hansen, 2009). The consequence of this was that it made the time for testing briefer than when trials to error-free performance was used and it also provided a statistically defensible basis for standardized task administration.

In the late 1980’s, the Austin HPMLT was used by some researchers to study impairments in spatial mapping in clinical conditions proposed to be associated with disruption to the hippocampus, such as alcohol dependence (Bowden, 1988, 1989;
Bowden & McCarter, 1993; O’Brien, Bowden, Bardenhagen, & Cook, 2003). The rapid encoding and recall of object-locations has consistently been shown to depend on the right mesial temporal region (Saling, 2009). However, the neurological impact of long-term chronic alcohol use on brain structures associated with executive functions has not, to our knowledge, been assessed using error analysis in the Austin HPMLT. The Austin HPMLT was used extensively in clinical neuropsychological studies, including normative and validation studies in healthy adults (e.g. Bowden & Smith, 1994; Bowden, et al., 1992; Crowe et al., 1999; Morrison & Gates, 1988; Tucker, Kinsella, Gawith, & Harrison, 1987). Studies that administered the Austin HPMLT in neurologically impaired populations, in which executive functions would likely be affected, typically define impairment in terms of the total errors, total trials, and total time to reach some criterion (Darby & Walsh, 2005; Österberg et al., 2009; Österberg, Österberg, Órbæk, & Karlson, 2002; Órbæk, Karlson, Bergendorf, & Seger, 2000; Barker et al., 2005; Kinsella et al., 1995). Despite the extensive use of the Austin HPMLT it is difficult to integrate data across studies because no standard method for administration and scoring of the task has been developed or validated. For example, studies using the Austin HPMLT differ in the criterion for task completion (10 trials: Barker et al., 2005; Bowden et al., 1997; O’Brien et al., 2003; Österberg et al., 2009; Walton & Bowden, 1997; two, or three error free trials: Kilpatrick et al., 1997; Kinsella, et al., 1995; Kinsella et al., 1997; Mathais & Kent, 1998; Morrison & Gates, 1988). More importantly, although the task rules have been described by Bowden and by Milner (Bowden, 1988, 1989; Milner, 1965), studies using the Austin HPMLT have been inconsistent in reporting which, or if any rules were given to participants prior to assessment (Barker et al., 2005; Matias & Kent,
As described earlier, Bowden & Smith (1994) described the wide variance in trials to criterion across control and clinical groups in the Austin HPMLT. Furthermore, trials to criterion and errors to criterion are also highly correlated in healthy adults \((r = .73-.91: \text{Tucker et al. 1987})\). Bowden has also shown that test-retest reliability on alternative pathways of the Austin HPMLT is generally low (errors to criterion: \(r = .77\), trials to criterion: \(r = .64: \text{Bowden, 1989}\)). The rules of the Austin maze are given verbally and task administration does not contain training, or pre-trial practice. This raises the possibility that improved performance on the second administered pathway (e.g., Bowden, 1989) reflected acquisition of the task rules in the first administration that was consolidated by the second administration.

Searching the same location, or within-search errors were regarded by Milner (1965) as errors common to patients with frontal lobe injuries and interpreted as ‘perseveration’, but they are not included as rules as reported by any studies published using the Austin or Milner HPMLT. The definitions of rule break errors have also been inconsistent, when they have been recorded in the Milner HPMLT. Bowden and Smith (1994) criticized the use of rule break errors for making clinical decisions because of the absence of normative guidelines for determining impairment in this measure. However, there were at least 5 studies before the mid 1990’s that indicated the rates of rule break errors in different clinical groups was sufficiently high to warrant analyses as a separate measure to other error categories (Behar et al., 1984; approximately 50-to-60 in patients with frontal lobe lesions over 15 trials: Canavan, 1983; 20 rule break errors in patients with frontal lobe injuries on trials to criterion: Milner, 1965; approximately 2-6 rule break errors in teenagers with Trichotillomania,
obsessive compulsive disorder, or Mixed anxiety disorders in 10 trials: Rettew et al., 1991; approximately 42-to32 rule break errors in younger and older schizophrenics respectively, in trials to criterion: Steinberg et al., 1983). By comparison, studies that included healthy controls showed that rule break errors were infrequent in this group and ranged from approximately 1-to 5 errors in trials to criterion (17 trials: Milner, 1965; 20 trials: Steinberg et al., 1983: 10 trials: Rettew et al., 1991). Even when rule break errors have been reported, the theoretical significance of these errors in the stepping-stone HPMLT has received little attention. Subsequent studies using the Austin HPMLT have not considered the quantitative and qualitative nature of errors from which a theoretical analysis of task variability could be advanced.

Although the Austin HPMLT has been used widely in clinical neuropsychological contexts, this review suggests that it still remains more of an experimental than clinical neuropsychological test. Researchers have sought to characterize variability in the rate of pathway learning errors in order to understand the metric properties of the task, as much as to investigate cognition in different groups. Furthermore the construction of the Austin HPMLT, as a commercial manual task without alternate pathways, limited the administration of the task to verbal instructions without practice or task familiarization. In order to use HPMLTs for the study of cognitive functions, as opposed to being a tool for understanding the nature of maze learning, recent computerized adaptations of the HPMLT paradigm have become more standardized, with rules to guide performance and criteria for classification of errors specified. Two examples of such adaptations are discussed below.

3.1. Computerized versions of HPMLT.
Computerization of the maze format from the late 1980’s had several advantages for administering and analyzing performance on HPMLTs (Morrison & Gates, 1988). First, computerization allowed greater flexibility of HPMLT design and administration. Second, computerization makes it possible to record and classify automatically every move, as well as every move sequence with respect to prior moves or to the maze layout. Functional neuroimaging studies using HPMLT have shown changes from frontal to temporoparietal activity over practiced trials, consistent with a shift from trial-and-error searching when the pathway is unknown, to navigation from memory (Van Horn et al., 1998). This analysis is also consistent with the interpretation of HPMLTs as a spatial learning paradigm within the context of executive functions (Milner, 1965). At least two versions of this HPMLT were designed in the first decade of 2000 for use in clinical research, and imaging environments respectively. This section of the review begins with an examination of the neurological basis of stepping-stone maze performance in healthy adults from a neuro-imaging study, followed by consideration of computerized HPMLTs. The Groton Maze Learning Task (Schroder et al., 2004) preceded the development of the Integneuro maze (Paul et al., 2005). However, in order to critically examine common limitations of HPMLT studies in general, the Integneuro maze will be examined before consideration of studies using the Groton Maze Learning Task.

Hidden pathway maze learning has been described as a task involving mental mapping (Bowden, 1989), and has been associated with topographical memory (De Renzi et al., 1977; Habib & Sirigu, 1987). However topographical amnesia has been defined as environmental disorientation in path finding involving locomotor movement through spaces that are beyond the immediate perceptual horizon (Habib & Sirigu, 1987). It has not been established that the stepping stone maze
reflects the same basic processes of way-finding, or navigation in real or virtual landscapes. For example, in hidden pathway maze learning the entire problem space is visible (Barker, 1931). Hence it is not necessary to create a mental map of an unseen topology. Milner (1965) suggested that the frame of the HPMLT task was fixed in relation to the body and therefore does not require the transformation of coordinate reference frames that is required by locomotor tasks. It is unlikely that pathway mapping relies exclusively on visual imagery: One study has shown that attempts to draw the pathway in conjunction with manual maze learning conferred no advantage to pathway encoding (Mann & Jewell, 1941). Another study has shown that there was no transfer of learning between an identical stylus and a physical maze (Jones & Batalla, 1944). It is therefore probable that maze performance relies on exocentric spatial representations that are dependant on the hippocampus and supporting regions rather than exclusively visual processes.

Although there have been neuro-imaging studies of hidden pathway maze learning, the tasks used have differed to the standard Milner versions (i.e. do not allow for most rule break errors), have examined non-specific regional measures of brain activity or connectivity, or have not examined the performance of healthy adults (Chen, Chauh, Sim, & Chee, 2009; Clark et al., 2006; Flitman, O’Grady, Cooper, & Grafman, 1997; Paul et al., 2009; Tate et al., 2010). Only one study has examined stepping-stone maze performance in healthy adults using positron emission tomography (PET).

Van Horn et al., (1998) used PET imaging to examine the neural correlates of novel, and practiced maze learning using a computerized 36-step maze. A chase test was used as control. Initial maze performance elicited activity in the right prefrontal cortex (Brodman’s areas (BA) 8, 9, 45, and 46). Minor activity was
observed in anterior cingulated and medial prefrontal area’s (24, and 32), left middle frontal gyrus (10), right inferior parietal lobule (7, 40), left lingual gyrus (19), and cuneus (18, 19). Practiced trials (2.5 minutes) elicited strongest parietal activity (bilateral medial parietal and precuneus (7, 31), and posterior cingulate, or retrosplenial cortex (29, 30). Little activity was observed in the prefrontal cortex in practiced trials, but activity was observed in left anterior prefrontal cortex in initial trials (10). Together the pattern of activity showed initial right-sided anterior activity during task acquisition, along with BA 34, which has shown to be involved in spatial encoding (Chrastil, 2013; Van Horn et al., 1998), and medial/posterior activity with recall. Hippocampal and parahippocampal areas were active during both encoding and recall, suggesting the involvement of the place navigation system (O’Keefe & Nadel, 1998) and the grid system in the entorhinal complex for recall. Brain areas activated in initial and practiced trials are shown in Figures 3.1.1 and 3.1.2. The functional significance of the regions of interest from this study can be explored in greater detail in Brodmann’s Interactive Atlas: [http://www.fmriconsulting.com/brodmann/Introduction.html](http://www.fmriconsulting.com/brodmann/Introduction.html)

FIGURE 3.1.1 ABOUT HERE

FIGURE 3.1.2 ABOUT HERE

It must be noted that responses in the Van Horn et al., (1998) study was via a hand-held device, presumably a joystick or a four-button keypad. If so, then only individual horizontal and vertical steps were possible, thus reducing the rule structure of the task and response requirements from the original Milner maze.
Therefore the results from this study should be interpreted with caution in relation to the 28-step manual stepping-stone maze format.

3.2. The Integneuro Maze.

The IntegNeuro battery from the Brain Resource Company has been used mostly for investigating the brain correlates of cognitive deficits associated with heterogeneous brain disorders (Lipton et al., 2009; Paul et al., 2009; Tate et al., 2010). The Integneuro maze is an 8 x 8 HPMLT presented on a touch screen. Moves are made using a 4-directional control panel (with arrows) set on the screen below the maze. Feedback is given by auditory and visual responses below the maze (not on each selected tile). Diagonal and jump moves are not possible.

The Integneuro maze task runs until two error-free trials, or a timeout of anywhere between 7 minutes (Clark et al., 2006; Mathersul, et al., 2009; Schofield et al., 2009) and 10 minutes (Gunstad et al., 2007; Paul et al., 2005; Paul et al., 2006; Paul et al., 2009; Silverstein et al., 2007). The computer software records the number of trials, errors, and time to completion (Clark et al., 2006). Overruns are also recorded, where participants continue to search locations in a linear trajectory when the pathway turns. The reason for inclusion of this measure and its theoretical significance has not been made explicit in any study that has used the Integneuro maze. Overruns have been reported previously and described as pattern factors of the maze that generalize to a local forward drift momentum, or orientation direction in choosing pathway locations (Jones & Yoshioka, 1938). In other words, if the pathway is unknown, participants tend to continue searching in the same goal-ward direction until an error is indicated. The pathway consists of 24 steps (Integneuro User Manual Version 2007v3: available from www.brainclinics.com), however, the
number of turns included in these steps has not been reported and there is no evidence as to whether different pathways are used. The Integneuro maze is described as a test of executive functions based on the Austin HPMLT, and there is extensive normative data from 6-years to 70+ in 1007 participants (Clark et al., 2006), and one validation study (Paul et al., 2005). However this validation study did not use any other form of maze task for comparison (for example the Austin maze on which the task was purportedly based: Paul et al., 2005).

The Integneuro maze appears to be based on the Morrison and Gates (1988) computerized version of the Austin maze in that moves are executed via a 4-directional keypad made remotely from the maze grid, not within the grid itself. The move restrictions of this format are similar to the physical ‘alleys and stops’ of the Oxford maze and its predecessors in which diagonal or jump moves cannot be made (Jones & Joshioka, 1938). However because move decisions are executed remotely to the maze environment, it is possible that move sequences could be coded linguistically according to application of a sequence of commands restricted to ‘up, down, left, right’. Because these commands can be applied and evaluated without visual inspection of the maze layout it is possible that the Integneuro maze task could be completed efficiently from verbal memory alone, or that any visual decision making could be supported or enhanced by linguistic processing. Therefore differences in the rule structure and administration of the Integneuro maze limits comparison to models of HPMLT maze learning based on performance on the Austin and Milner HPMLT as measures of either executive functions or spatial mapping.

In one study, the correlation between Integneuro maze completion time and the Rey Complex Figure Test-delay (RCFT-d) was taken as a validation of the Integneuro maze (Paul et al., 2005). However the selection for the RCFT measure in the delayed
form as a validating task is unclear. While the authors assert that the Integneuro maze is a measure of executive functions, the delayed trial of the RCFT is considered to be a task of visual memory (Caffara, Vezzadini, Dieci, Zonato, & Venneri, 2002). The use of the RCFT as a test of executive function depends on the analysis of error patterns, and organization abilities (Meyers & Meyers, 1995; Meyers & Volbrecht, 1998). The common operations required by the Integneuro maze and the RTDFT-d were not discussed. Divergent validity was established by the lack of correlation between Inegneuro maze measures and the California Verbal Learning Test- Research (short delay). However at least 30 measures were recorded from 12 tests administered in the experiment, but an inter-correlation matrix of other measures in the battery was not reported. The veracity of the claim of validation of the Integneuro maze cannot be assessed without a more complete report of the convergent and divergent relationships with other tests of cognitive functions administered in the experiment. No such data has as yet been published for this version of the HPMLT.

The Integneuro maze has been used extensively in imaging and tractography studies as a task of executive function within the context of spatial learning (Lipton et al., 2009; Schofield et al., 2009; Seckfort et al., 2008; Tate et al., 2010). However imaging studies using the Integneuro maze have used summary measures over trials, not behavioral trial by trial data in relation to stability or changes in brain functions over trials, as was shown by Van Horn et al., (1998). The contribution of studies using the Integneuro maze for understanding HPMLT performance is limited, because the maze has not been validated as a HPMLT, and the theoretical significance of the measures it records, such as overruns are not well understood.

3.3. The Groton Maze Learning Task (GMLT).
The Groton Maze Learning Test (GMLT) is a touch screen HPMLT, based on the Austin and Milner HPMLT (see Figure 3.3.1). Originally developed by P.J. Snyder and colleagues for use in neuropharmacological studies at the former main campus for research and development of Pfizer Inc. (Boulanger, Snyder, & Cohen, 2006; Castner et al., 2004; Pietrzak et al., 2007). Maze pathways begin at the top left-hand corner and end at the bottom right-hand corner (see Figure 3.3.1). The pathway remains hidden and only the most recently selected square is illuminated. When a correct square is touched, it turns white with a green tick, accompanied by an auditory tone. When an incorrect tile is touched, it turns white with a red cross, and no tone is emitted. Move feedback appears for 250 ms, following which, the square changes color to blue. If two successive errors are made, the last correct location (referred to as the “head of path”) flashes 10 times with a tick, as a signal to return to the head of path. Correct and error-move feedback provides pathway information (Pietrzak et al., 2009a; Pietrzak et al., 2009b). For the GMLT, the grid size, pathway length/number of turns, and rules are the same as for the Milner HPMLT.

Because the GMLT was specifically designed for clinical trials, there was emphasis placed on the ability to generate alternative forms that were resistant to practice effects (Maruff et al., 2006; Pietrzak et al., 2009b; Snyder et al., 2005a; Snyder et al., 2005b; Snyder, et al., 2008a). Hence, there are 20 alternate-form maze pathways in the GMLT, and these are all equivalent with respect to the number of correct moves, turns, and individual choice or decision points, within
the 10 x 10 grid (Fig. 3.3.1). Each time an individual starts the test, a form is
selected in pseudo-random order to ensure that no individual subject will see the
same alternate form within three successive test sessions. When a subject attempts
to discover the hidden maze on the first trial, selection of tiles that are allowed by
the task rules for finding the hidden pathway are categorized by the GMLT
software as exploratory errors. Moves that violate the task rules are classified as
rule break errors (see Table 3.3.1 for a description of error categories).
Researchers using the GMLT were the first to systematically codify error
categories in HPMLTs, as well as the first to report analysis of different errors as a
standard output feature (e.g. Maruff et al., 2006; Snyder et al., 2005a; Snyder et al.,
2005b). Each step in the pathway must be located in sequence, according to the
rules in order to complete a trial.

| TABLE 3.3.1 ABOUT HERE |

From any location in the GMLT, it is possible to move to any location within the
grid. For the GMLT software to respond with appropriate feedback to every move, it
is necessary that all moves can be classified according to the task rules and according
to the nature of the move (i.e. see Table 3.3.1 for error classifications). A
consequence of this is that the GMLT output can be rendered to provide raw data for
each move and each response time, as well as trial-by-trial summaries of different
errors and responses. Because the GMLT records and classifies every move, there are
more categories of ‘erroneous’ moves than are reported by other HPMLTs (double
tapping tiles, moves between tiles). Correctly returning to the head of path after
errors are signaled can also provide a measure of error feedback use, although this
measure has not been previously reported. Most reports using the GMLT include
incorrect moves that comply with the task rules (legal errors); errors that do not
comply with the task rules (rule break errors); and repeating the same error in succession (perseverative errors). These error categories are typically summed over the five learning trials and expressed as an effect size of difference from control, or change from pre-treatment baseline score (Pietrzak et al., 2008a; Snyder et al., 2005a; Snyder et al., 2008a; Thomas et al., 2008). Practice trials are an essential component of the GMLT, as it is important to ensure that participants fully understand the rules of the task and can demonstrate stable baseline performance prior to administering the task in research or clinical applications. Other studies have reported a summary measure of learning efficiency, or a ratio of total moves over trials to task completion (Collie, Maruff, Snyder, Darekar, & Huggins, 2006; Maruff, et al., 2006; Pietrzak et al., 2007; Pietrzak et al., 2009a; Pietrzak, Sprague, & Snyder, 2008b; Snyder et al., 2005a). There is ample evidence in the references above suggesting the sensitivity of GMLT measures to detect clinically significant change or difference using standard outcome measures of error and efficiency. However, these along with other GMLT measures can be used to address a variety of theoretical questions related to the specificity and qualitative nature of cognitive performance on a complex task.

In neuropsychological research and practice, it is important that the choice of outcome measures is suitable for specific experimental and clinical questions. For example, summary outcome measures may be preferred to multiple, “process-based” measures when a test is included as part of a large test battery, or when non-specific brain disease or disorders would likely be associated with diffuse cognitive impairments, such as in schizophrenia (Pietrzak et al., 2009b). Alternatively, researchers may be interested in the component operations underlying task performance in which multiple measures are useful (e.g. Alderman, Burgess, Knight, & Henman, 2003). As the GMLT is a learning task, it is also possible to examine the
learning process through the change in different parameters of performance over trials, although few studies have examined individual trial data in the GMLT (Pietrzak, et al., 2009a; Thomas et al., 2013).

One of the most consistent findings in studies using the GMLT is the differential effect of clinical status on measures of rule break, and legal or exploratory errors. Although there is vast amount of evidence from early studies confirming that HPMLTs measure some aspects of executive functions in addition to spatial mapping and learning functions, these processes can be differentiated on the GMLT. For example, exploratory and rule break errors can be independently affected in clinical groups, as well as with pharmacological manipulation of neurotransmitter systems. Memory and executive function deficits often co-occur in cases of diffuse brain disorders and several studies have shown the dissociation of rule use and spatial memory processes in hidden pathway maze learning (Canavan, 1983; Thomas et al., 2008; Snyder et al., 2005a; Snyder et al., 2008b). Rule-break errors in HPMLTs are rare in healthy adults, but common in populations with executive function deficits due to anatomical, chemical, or age-related disruption to the prefrontal cortex (Pietrzak et al., 2008a; Snyder et al., 2008a; Snyder et al., 2008b; Thomas et al., 2008). In these studies, spatial memory in the GMLT is generally far less affected than executive functions (e.g. Snyder et al., 2008a; 2008b; Thomas et al., 2008). By contrast, patients with focal lesions of the hippocampus, reduced performance on HPMLTs are predominantly driven by reductions in spatial learning and memory, with little or no disruption to rule use ability (Milner, 1965; Milner et al., 1968). Importantly, the ability to understand the relative contribution of cognitive systems to performance can inform neuropsychological models of hidden pathway maze learning. For example, Snyder et al. (2008b) found that children with attention deficit hyperactivity disorder
(ADHD) made more rule break errors than age-matched controls, but not more exploratory errors in the GMLT. Administration of stimulant medication in children with ADHD led to a greater magnitude reduction in legal than perseverative errors, suggesting that stimulant medication improves both learning/memory and executive processes in children with ADHD to differing degree’s. Together, these findings are theoretically and clinically important because they suggest that it is possible to isolate and compare different task components of HPMLTs that allows for greater interpretive scope in analyzing performance variability.

Because every move in HPMLTs represents a discrete decision, behavioral analysis of move-by-move performance can be coupled with neuroimaging data. Although HPMLTs, other than the Integneuro maze, have not been widely used in imaging studies, they show great potential for coupling simple responses with brain activity and cognitive behavior in a complex task. For example, by examining the brain responses to individual moves, such as returning after errors as opposed to failing to return after errors, it is possible to make theoretical claims regarding the brain-behavior relationship in error monitoring, although such an analysis has yet to be conducted.

Mathewson et al. (2007) investigated error-feedback processing in aging by examining the event related potentials to error feedback in older and younger adults performing three levels of maze difficulty. Fifteen younger females (18-26 years), and 15 older females (65-87 years) competed 8 x 4 x 4 maze (1 learn, 2 test trials), 46 x 6 mazes (1 learn, 3 test trials), and 28 x 8 mazes (1 learn, 5 test trials). The main finding was that in younger adults, electroencephalogram (EEG) feedback related negativity was sensitive to errors at all levels of difficulty in both learn and test trials (it should be noted that Mathewson et al., did not differentiate error types).
Consistent with the findings of Van Horn et al. (1998) the maxima of EEG activity was evident in medial brain areas, including the anterior cingulate cortex (ACC), with an anterior shift evident in more difficult learn trials, and posterior shift evident in practice trials. For older adults, there was no specific response differentiation in the medial prefrontal cortex; neural responses to correct moves and errors were associated with more diffuse brain activity. Behavioral data also showed that there were no differences in errors on learning trials at each level of difficulty, but that older adults were less efficient in learning the mazes. Processing speed (Trails A) did not account for the variance in error rates, suggesting that performance was dependent on higher-order learning and executive processes. Older adults also produced smaller feedback-related-negativity (FRN’s) associated with error feedback in all conditions, suggesting that older adults may have recruited brain regions less reliant on dopaminergic activity (i.e. than the ACC). For both groups, the P3 was sensitive to feedback valence. The differentiation between event-related-potentials (ERP) and P3 sensitivity to error feedback (correct and incorrect moves) with age also support the inference that these two systems may rely on different neurotransmitter systems. One confounding factor in this study was that neural responses were not analyzed separately for qualitatively different error types (rule break and exploratory). Based on developmental studies of maze performance, it would be expected that rule breaking errors would be more frequent in older adults than in younger adults (Pietrzak et al., 2007). Analysis of qualitatively different errors could provide information regarding the brain-behavior relationship between internally generated errors (rule breaking), feedback errors (failing to return, within search), and correct error discrimination (returning after an error). It is likely that qualitatively different errors would be associated with distinct patterns of brain activity. If so, then
aggregating errors (rule break and legal) would appear as more diffuse activity than consideration of specific categories of error. This could account for the more diffuse error-related neural responses observed in the older adult sample. If this interpretation is correct, it underscores the importance of assessing different error responses in making inferences regarding putative brain bases of error monitoring processes.

As noted earlier in this review, Walsh (1985) hypothesized that the main difficulty underlying rule break errors in clinical populations with injuries to the prefrontal cortex was an inability to use error information to plan and execute moves, or error utilization. Rule break errors can occur as a failure to follow rules, such as ‘jumps,’ diagonals, or backtracking, or as a failure to use prior task information to select the next move, such as failing to return after error signals are given, or repeatedly searching the same location within a trial. Error analysis of these different types of errors could be used to identify the nature of rule use difficulties in normal development, as well as in clinical samples. Such an analysis would thus provide useful information regarding the theoretical basis of variability in maze learning.

An example of componential analysis can be seen in one study in which the GMLT was administered to children from 5-to-9-years of age (Thomas et al., 2011). Error analysis in this study showed an age-related trend in the profile of different error categories (Thomas et al., 2011). Compared to older children, younger children made more errors related to the ability to use error information (i.e., failing to return and within search errors), than failing to follow the rules (i.e., making ‘jumps’ and diagonal moves). This finding is consistent with the early development of rule use, but the relatively latter development of error monitoring in childhood (Bunge & Crone, 2009; Marchovitch & Zelazo, 2009). It is also consistent with Walsh’s (1985)
hypothesis that error utilization is the main difficulty demonstrated by individuals with executive function deficits in performance on the Austin Maze.

Computerization of HPMLTs have allowed for fine-grained analysis of component operations necessary for goal-directed behavior in trial-and-error learning, such as correct and incorrect error feedback use, spatial learning, and rule use. As such, this paradigm shows great potential as a tool from which models of complex cognition and its disturbances can be developed, based on the qualitative and quantitative nature of performance on a single test that is brief to administer, simple to perform, and resistant to practice effects.

4.1. Summary and directions for future research.

Maze learning was a central methodology for studying fundamental aspects of cognition and behavior throughout the 20th century. HPMLTs were designed to overcome the limitations of other maze formats in humans, and to improve the psychometric properties of such tests in order to augment interpretation of performance. With the advent of brain-based approaches to adult neuropsychology in the 1950’s, maze learning became a useful diagnostic measure of mid-temporal lobe and frontal lobe function according to the profile of different errors made by patient groups. Theoretical accounts of what HPMLTs measured during the 1980’s and 1990’s were restricted by the use of summary outcome measures, such as total errors or trials to a criterion of task completion. The use of global measures and lack of task standardization have consequently prevented comprehensive analysis of brain-based, or cognitive systems that could account for performance variability both within, and across studies. However, on the weight of evidence, modern HPMLTs are able to assess various cognitive abilities sub-served by functionally separate neural systems for spatial learning/memory, and goal-
directed error monitoring and rule use. The ability to measure different aspects of performance, as well as the potential to modify the task makes the HPMLT paradigm a powerful tool for investigating how different cognitive functions change, or remain stable as a function of age, clinical status, or task demands. The flexibility of HPMLTs, in terms of the range of outcome measures derived from a move-by-move, and trial-by-trial basis allows for a comprehensive approach to characterizing performance parameters reflecting component aspects of visuospatial learning and executive function.

HPMLTs are learning tasks that are simpler to administer and measure than large-scale or virtual navigation tasks, and appear to elicit neural responses in similar brain regions to large-scale tasks. HPMLT would therefore be useful for investigating factors that facilitate, or disrupt spatial encoding and recall processes in healthy adults and children, and how these relate to ‘real-world’ spatial processing, or specific executive function disruptions (error feedback use, rule following, goal maintenance) in patients with focal or diffuse brain disorders.

The review presented here has shown that there is a sound neuropsychological basis within which performance on the HPMLT can be understood. This now provides a firm foundation for exploitation of HPMLTs in the future to assist in understanding disorders of the central nervous system that are characterised by impairment in executive function but where the neuropsychological substrate of this impairment in now known. For example, GMLT measures are sensitive to cognitive impairment in schizophrenic patients compared to controls, as well as treatment improvement and stability over time (Snyder et al., 2008a). Another aspect of HPMLT that bodes well for understanding complex executive functions is that this paradigm performs well in psychopharmacological studies. Repeated performance is
not associated with practice effects, therefore HPMLTs can be used to assess the effects of compounds known to influence executive functions and spatial memory, such as dopamine antagonists or nicotinic agonists. Finally, one notable absence in the literature of HPMLTs is an understanding of the functional neuroanatomy of the different error types. Future studies could combine componential analysis of different task parameters with experimental methods to further delineate the functional dynamics of HPMLT performance, such as imaging, TMS, as well as pharmacological manipulations.
References


description


Small, W. S. (1901). Experimental study of the mental processes of the rat. II. *The


Fig. 4

Fig. 5
Fig. 5. Plan of maze. The black line indicates the correct path.
Figure 1.3.1. Jensen, M. B. (1934). The effect of punishment by electric shock for errors on a raised finger maze. The blindfolded participant negotiates a maze with a stylus in the right hand. He receives shocks for errors via immersion of left-hand fingers in a salt-bath attached to an electric circuit.

Figure 1.3.2. The two-storey duplicate maze (Miles, 1927). The participant must learn a pathway through a stylus maze on the lower section that is hidden from view. The maze on the top panel is an inverted version of the bottom panel.

Figure 1.3.3. Stylus grooves in the Carr maze (Carr, 1921). Slotted grooves are used to hold the stylus in place within the maze from participant entry until completion of the maze.

Figure 1.3.4. The pathway of the Carr maze with the position of stops indicated by dashed lines across the pathway. Location ‘A’ represents the entry, and location ‘B’ represents the end position. See Griffith (1931) for an electrical Carr maze.

Figure 1.3.5. The standardization of maze alley length and number of blind alleys in a stylus maze used by Warden (1924a).

Figure 1.3.6. The electrical wiring diagram of the Barker (1931) maze. The bold line represents electrical circuits defining the pathway. Fine lines represent electrical circuits that signal errors (a buzzer), when an incorrect contiguous
location is selected that is not the pathway. Trial-and-error learning in the Barker Maze is defined by correct and incorrect searches to locations that surround the current position in the maze. It must be noted that the Barker Maze did not have specified rule structure.

Figure 1.3.7. The Mann and Jewell bolt head maze (1941). Participants were to find a pathway traversing the contra-lateral corners of the grid, one step at a time, returning after errors before resuming a search.

Figure 1.4.1. The Jones and Yoshioka stylus maze (1938).

Figure 1.4.2. The underside wiring diagram of the Jones & Yoshioka (1938) maze. Feedback for each selected location in the maze was given by circuits A (shaded), which activated a light, and B (unshaded), which activated a buzzer. Stops in the pathway were created by the use of brass strips (C and D) that closed the circuit and activated a signal to indicate an incorrect location.

Figure 1.4.3. Pathway patterns used in Jones and Yoshioka (1938) are shown in bold with the entry indicated at ‘A’.

Figure 2.3.1. The pathway of the Milner HPMLT (Milner, 1965).

Figure 2.3.2 (1) and 2.3.2 (2). The Oxford Stylus Maze and pathways (from Newcombe & Russell, 1969).
Figure 2.3.3. The Austin HPMLT. (Image reproduced with permission of NCounters Australia). Contralateral red buttons control the timer for the beginning and end of the maze. Red and green lights are activated at the top of the maze to indicate correct and incorrect choices. The NCounter maze has a choice of two pathways.

Figure 3.1.1. PET activity in Brodmans regions in initial maze learning on a modified Milner maze (+ -major activity, - minor activity in Brodmann’s area): +8: Supplementary motor, +9: Middle frontal gyrus (prefrontal cortex), +45: Inferior frontal gyrus Pars triangularis), +46: Anterior middle frontal gyrus (prefrontal cortex), -24: Anterior cingulate gyrus, -32: Anterior cingulate gyrus, -10: Middle frontal gyrus (prefrontal cortex), -7: Superior parietal lobule (secondary sensorimotor cortex), -40 Inferior parietal lobule (supramarginal gyrus), -19 Inferior occipital gyrus (secondary visual cortex), -18: Middle occipital gyrus (secondary visual cortex).

Figure 3.1.2. PET activity in Brodmans regions in practiced maze learning on a modified Milner maze: medial parietal and precuneus 7, 31, and posterior cingulate 29, 30.

Figure 3.3.1. The GMLT. The grid shows the current location (blue tile) and the end goal (red target). The pathway begins in the top left corner and finishes in the bottom right corner.
Table 3.3.1

*Definition of Errors on the Groton Maze Learning Task. All Measures are Recorded as Counts on Each Trial (Five Trials in Total).*

<table>
<thead>
<tr>
<th>Error type</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploratory</td>
<td>Moves in accordance with task rules.</td>
</tr>
<tr>
<td>Failure to return</td>
<td>Failing to return to the last correct location after an incorrect move.</td>
</tr>
<tr>
<td>Perseverative</td>
<td>The same error repeated after returning to the head of path.</td>
</tr>
<tr>
<td>Back</td>
<td>Moving backwards on the pathway.</td>
</tr>
<tr>
<td>Jump</td>
<td>Moving more than one tile from the current correct location.</td>
</tr>
<tr>
<td>Diagonal</td>
<td>Moving to locations diagonal to the current correct location.</td>
</tr>
<tr>
<td>Double tap</td>
<td>Tapping the same incorrect location twice.</td>
</tr>
<tr>
<td>Same tile tap</td>
<td>Tapping the correct location twice.</td>
</tr>
<tr>
<td>Between</td>
<td>Moves between tiles.</td>
</tr>
</tbody>
</table>
Out

Moves out of grid.

Re-searching locations allowed by the task

Within

rules that have been identified as the incorrect location for the next step in the pathway.