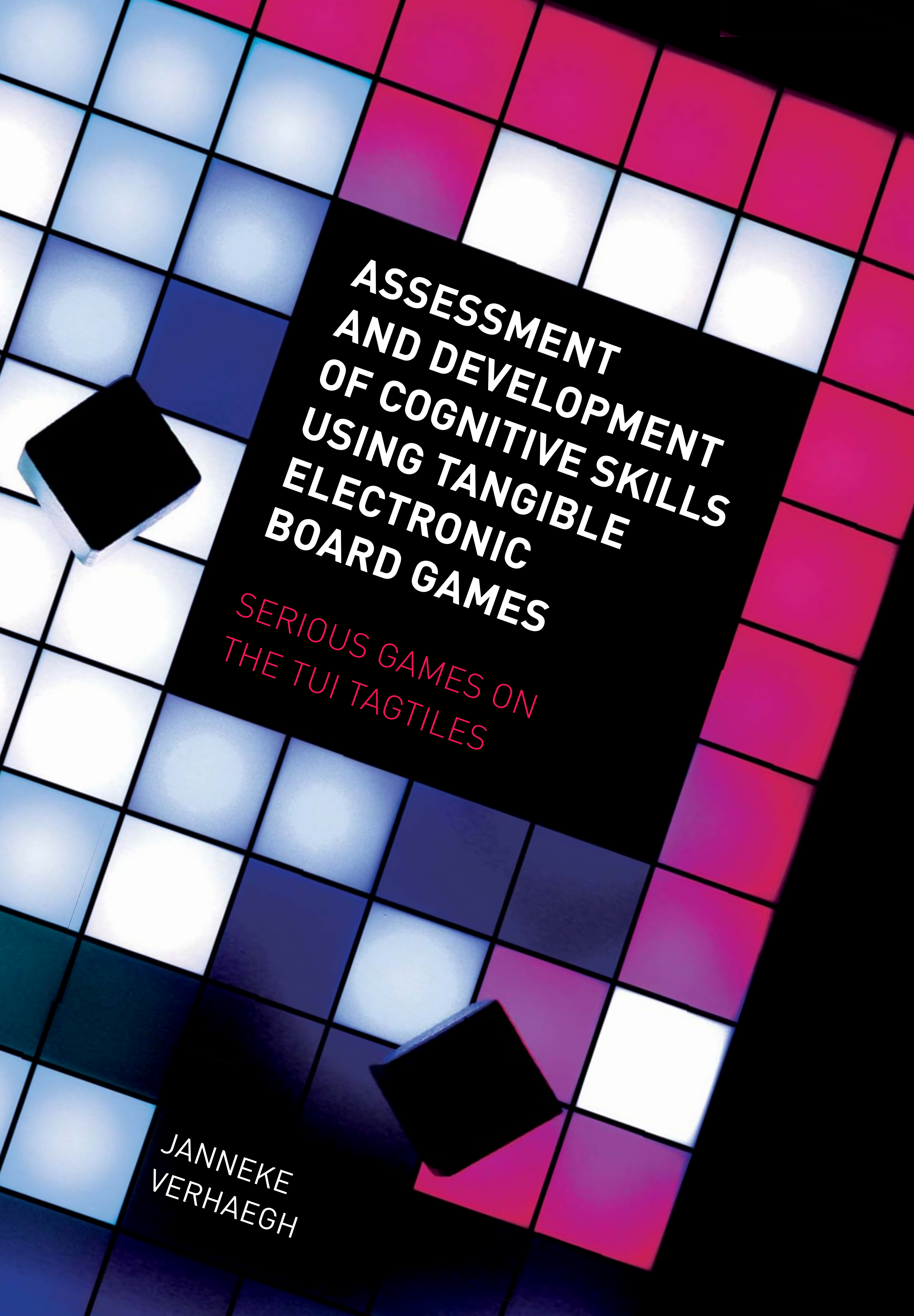
Assessment and development of cognitive skills using tangible electronic board games : serious games on the TUI TagTiles

DOI:
10.6100/IR739271

Document status and date:
Published: 01/01/2012

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)



ASSESSMENT AND DEVELOPMENT OF COGNITIVE SKILLS USING TANGIBLE ELECTRONIC BOARD GAMES

SERIOUS GAMES ON THE TUI TAGTILES

JANNEKE VERHAEGH

Assessment and development of cognitive skills using tangible electronic board games

Serious games on the TUI TagTiles

The work described in this thesis has been carried out at the Philips Research Laboratories in Eindhoven, The Netherlands, as part of the Philips Research programme.

Assessment and development of cognitive skills using tangible electronic board games, by Janneke Verhaegh
Eindhoven: Technische Universiteit Eindhoven, 2012. -Proefschrift-
A catalogue record is available from the Eindhoven University of Technology Library
ISBN: 978-90-386-3284-1

Cover design: Oranje Vormgevers
Lay-out: Grefo PrePress
Reproduction: TU/e Printservice

Assessment and development of cognitive skills using tangible electronic board games

Serious games on the TUI TagTiles

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
rector magnificus, prof.dr.ir. C.J. van Duijn, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op woensdag 28 november 2012 om 16.00 uur

door

Janneke Verhaegh

geboren te Wageningen

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr. E.H.L. Aarts
en
prof.dr. W.C.M. Resing

Copromotor:
dr.ir. W.F.J. Fontijn

Contents

1

Introduction

1.1 General introduction

Learning is the most important thing children do, next to growing. Parents would like their children to live a happy life and at the same time develop optimally. Learning in an enjoyable way is therefore a goal in itself. Research focusing on optimizing children's learning process is thus important for children and their parents, but also to society in general. Probably, children that develop optimally and exploit their talents will contribute most to society.

Play can be regarded as the most natural form of learning. Physical objects, whether or not designed as toys, are often at the centre of young children's play. They trigger exploration, like manipulating an object to test the effect, but they can also facilitate pretend play and shared play with other children. These explorations take place at the edge of children's capacities and often extend them. Play is therefore a crucial facilitator of children's development (e.g., Piaget, 1962; Goswami, 2008). Key is that play is a self-motivating, enjoyable activity, which offers a great way to learn new things in the real world (e.g., Abt, 1987; Salen & Zimmerman, 2004).

The role of play in learning has been first described by pioneers in the field of developmental psychology (e.g. Bruner, 1964; Piaget, 1962; Vygotsky, 1978) and has been a topic of interest ever since (e.g. Malone, 1980; Prensky 2002; Rieber, 1996). Games can be seen as a structured form of play (Salen & Zimmerman, 2004). This structure can be used to support learning through activities in a game. Games that are designed to have a primary purpose besides entertainment are referred to as 'serious games' (Abt, 1987).

Modern serious games are typically screen-based, or use 'virtual reality' as a playing environment. A different type of computerized interfaces, as a means to facilitate learning, has gained interest as well. Tangible User Interfaces (TUIs) are digitally enhanced physical objects, in which the visibility of the computer is reduced as much as possible (O'Malley & Stanton Fraser, 2004). The combination of TUIs and serious games has opened up a new field of research for learning.

TUIs are a part of a general development in technology where computing power is becoming available everywhere and in everything, as was predicted and strived for in visions like Ubiquitous Computing (Weiser, 1991), Pervasive Computing (Satyanarayanan, 2001) and Ambient Intelligence (Aarts & Marzano, 2003). These visions have started to become reality, including learning environments such as schools. The most obvious augmentation of the classroom environment is the integration of electronics and computing. The resulting options for interactivity and connectivity between systems, services and applications offer many opportunities to enrich the learning experience and amplify the educational effectiveness (Verhaegh, Fontijn, Aarts, Boer & Van de Wouw, 2011). Care should be taken, however, that the resulting situation fits the learners and their needs, and that a 'technology push' is avoided.

In this context, two aspects are of prime importance. The first aspect relates to the individual child and its developing abilities. For one, the electronics should be very accessible and easy to use by young children. The interfaces used should be simple and make sense to the children and the interaction with them should come natural to the child. This implies that children should not have to adapt to the new technology, but rather that the technology is adapted to fit the children. This also holds for the applications. As the child's skills develop, the application should adapt actively to the changing abilities (Verhaegh et al., 2011). Individual needs of the children related to learning styles and personal interests should be taken into account as well.

The second aspect relates to educational effect. Electronics and computation should only be used in classrooms when they can actually improve the current educational process. This may sound obvious, but too often the technology takes center stage and the benefits for education are simply assumed. The following objectives could be targeted by using such systems: promoting the active participation of each individual pupil, enhancing the possibility to share and collaborate with peers and teachers, motivating children for tasks that are currently less attractive and supporting the different needs of each individual, including early signaling and remedying shortfalls in development. Direct feedback and assistance from the system can be part of this. Taking both aspects into account should result in applications that are at the same time intrinsically motivating and educationally effective (Verhaegh et al., 2011).

Besides the options that TUIs provide for education, the interaction that they enable provides opportunities in a more general context as well. The devices that people are surrounded by are becoming increasingly complex. At the same time, there is less time available and less willingness to read the manuals that may accompany a newly bought device. Therefore a transition is needed from the situation where people need to learn to understand a device, to the reversed: the device needs to learn to understand the people who control it. A rather novel concept in this context is that of implicit control (Fontijn, 2007) of a device: the device 'observes' the human behavior and adapts itself, to be able to increasingly understand the wishes of the user, and anticipate these. Also the common explicit control that people use to control devices needs to change. For example a television has much more functionalities than in its early days and cannot be simply controlled through physical buttons only. To support the user, virtual context information can be provided to simplify the interaction. Controlling devices can be made easier by making them easy to learn to use, and also enable interaction that comes natural to people (Fontijn, 2007), preferably in a manner that is enjoyable to the user. Since the manipulation of physical objects and interfaces comes natural to people already at a very young age (e.g. Piaget, 1962), tangible interfaces seem to be a good candidate to support enjoyable, playful interaction.

1.2 Research focus

There are a number of challenges to developing effective educational TUIs. First, how can it be proven that a TUI is effective for learning? This requires availability of appropriate hardware and application software or content. An existing TUI, the TagTiles console, was used as a testing platform for the studies in this dissertation. Thus, assuming that the hardware is available by using the TagTiles console, then how can appropriate software be created?

A number of requirements, based on the theory described in Chapter 2, were defined for creating the content: (1) the content should be adaptive to the progress in the child's development and (2) it should be self-motivating; (3) it should be designed such that (a) the emphasis lies on training the intended skills, avoiding distraction by other skills that may be needed in the task. To this end, (b) the content development process should involve domain experts such as educational experts, or ideally, they are enabled to make their own software*.

The general hypothesis that is addressed in this dissertation is:

We can create games for a tangible user interface that are effective for developing the cognitive skills of children.

With 'effective' it is meant that the game is self-motivating, adaptive and with the proper educational gains, i.e., the child learns what is intended to be learned by the task-designer.

---

* The latter is left outside the scope of this dissertation (this is addressed elsewhere; see Van Herk, Verhaegh & Fontijn, 2009).

Whether specific (cognitive) skills can be effectively addressed with a self-motivating adaptive game is examined in three stages, targeting the following hypotheses:

(1) TUIs can be effective educational interfaces. (Chapter 3)
(2) Specific (cognitive) skills can be addressed. (Chapter 4)
(3) A self-motivating and adaptive game can be developed based on the outcome of hypotheses 1 and 2. (Chapter 5)

1.3 Outline

The theoretical background to this dissertation is presented in Chapter 2. Relevant theories of cognitive development, learning and motivation are described. Some background is provided to TUIs and it is explained why they have gained attention for educational use. Further, the TagTiles console and a number of prior experiments with it are described. Subsequently, the studies presented in the following chapters are introduced and also an approach to designing adaptive educational games is described.

In Chapter 3, a study is described in which the use of a virtual, PC-based interface was compared to the use of a tangible, non-electronic interface for the same puzzle task. With this study the influence of the type of interface on the performance of children on an educational task was investigated.

In Chapter 4, a study is presented in which the performances of children on the TagTiles tasks were correlated with their performances on conventional nonverbal intelligence (sub)tests. This study was carried out to validate whether a range of TagTiles tasks can be used to address nonverbal, cognitive skills.

Chapter 5 describes a study carried out with 'Tap the little hedgehog', a game based on the tasks tested in Chapter 4. The study establishes whether the tasks kept their ability to address certain nonverbal skills in the context of the game, whether children were able to play the game independently and whether they experienced it as an enjoyable, self-motivating game.

The summarized results of the empirical studies, possible applications of the work and the conclusions are provided in Chapter 6.

Chapters 3 and 4 have been published, and Chapter 5 submitted, as regular contributions to the scientific literature and these chapters are included without a change. As a result, some of the information in these chapters is redundant.

2

Theoretical background

In this chapter the theories and prior empirical studies relevant for this thesis are presented. Principles of cognitive development and learning are described, as well as the role of motivation in learning. Serious games are introduced as a learning environment in which a child's self-motivation can be utilized. Subsequently, TUIs and their use for educational applications are discussed. As an example of an educational TUI, the TagTiles console and its use for integral skill development are illustrated with a number of design cases that were conducted. Furthermore, an introduction is provided to the experiments in chapters 3, 4 and 5. Finally it is described how the implementation of scaffolding can guide the adaptation of the game play to the child's progress in serious games.

2.1 Theories of cognitive development, learning and motivation

2.1.1 Cognitive development

One of the most influential theories on cognitive development was described by Jean Piaget (1896-1980). He was interested in the nature and origins of knowledge (Bidell & Fisher, 1992). According to Piaget, infants learn by exploring objects around them, thereby actively constructing 'schemas'. The formation of schemas depends on the development of sensory-motor abilities, and these determine the ability to coordinate responses to a complex environment (Farnham-Diggory, 1992). Thus, cognitive development depends mostly on sensory-motor responses (Goswami, 2008). For example, a young infant explores the world by sucking on objects, and as such is able to classify objects as 'suckable', or 'non-suckable', and later on also as 'slappable' or 'bangable' or 'scratchable'. The formation of these concepts eventually enables understanding of more abstract concepts, such as time and space (e.g., Berk, 2008). Cognitive schemas are constantly adapted by a process in which children seek to find a balance between what they encounter in their environment and the cognitive processes and structures they bring to the encounter. In Piaget's terms, new knowledge that easily fits with old ideas will be assimilated. New knowledge that does not fit in prior knowledge causes a reorganization of ideas before it can be taken up and this process is called accommodation (Piaget, 1952). Piaget's view is referred to as 'constructivism'; children construct new schemes and develop new forms of organization (e.g., Morra, Gobbo, Marini & Sheese, 2008; Siegler, DeLoache & Eisenberg, 2011). Constructivism was controversial in Piaget's time, as the mainstream English-language tradition assumed a Cartesian model as a basis of intelligence in which the mind was assumed to be isolated from the body (even from the world) and thus, knowledge could not result from action (Bidell & Fisher, 1992).

Piaget's best known contribution to developmental theory is his stage theory of the progression of cognitive development (Bidell & Fisher, 1992). This theory is based on the assumption that learning and thinking develop in a stage-like fashion. The stages can be found in Table 2.1. As children pass through these stages, their cognitive structures become increasingly sophisticated and abstract. Piaget claimed that all normal children go through the same stages of development in the same order. These stages cannot be skipped or accelerated; not even by knowledge and skills acquired through training. In Piaget's view, major shifts in thinking take place at approximately the ages of 2, 7 and 12 (e.g., Slater, Hocking, & Loose, 2003).

This staged view of development has not remained undisputed (e.g., Siegler & Alibali, 2005). One of the main criticisms is that the stages Piaget proposed, do not account for the large variations in cognitive development that occur amongst individuals. Although some psychological skills do seem to follow this staged development, such as the occurrence of abstract thinking, for other skills it could never be confirmed. Although the notion of universal stages is the most well-known part of his theory, the core of his work concerned individual variability in cognitive development (Goswami, 2008). Furthermore, the ages that Piaget related to the stages are nowadays thought to be too high. Research has demonstrated that children show the described cognitive characteristics at a younger age than Piaget had claimed (e.g., Goswami, 2008). Also, the idea that action is the main source of cognitive development has been criticized. It was thought that knowledge was built through perception primarily (Bidell & Fisher, 1992). However, more recent insights have provided support for Piaget's account. As Goswami described, data available from neuro-imaging studies have indicated that it is very likely that sensory-motor behaviors play an important role in knowledge acquisition (Goswami, 2008).

Table 2.1. Piaget's stages of cognitive development derived from Siegler, DeLoache, & Eisenberg (2011).

Age	**Developmental stage**	**Cognitive characteristics**
Birth-2 2 years	Sensorimotor	Infants understand the world through sensory information and motor responses. (Major accomplishment: understanding that an object does not stop existing once hidden from view: object permanence.)
2-6 years	Preoperational	Children can use mental representation to reason about the world, but thinking is not yet logical. (Major accomplishment: language and conceptual development.)
7-11 years	Concrete operational	Children can perform logical operations on concrete objects. (Major accomplishment: understanding principles of conservation of quantity.)
>12 years	Formal operations	Children can think logically about abstract issues and hypothetical situations.

A second major theory about the construction of knowledge is Lev Vygotsky's (1896-1934) sociocultural theory of development (Vygotsky, 1978). According to Vygotsky, a child's development is determined by his or her social environment. Thus, an individual's higher mental processes have their origin in social processes, according to Vygotsky. This premise includes the idea that there are no single, distinguishable stages in development. Furthermore, Vygotsky believed that children's development is complex due to an ongoing reorganization of the process of development itself. Related to the importance that he attributed to the effect of a child's cultural and social environment on cognitive development, he emphasized the importance of language in cognitive development. He argued that the emergence of language in a child leads to new use of 'tools' and behavior. Also, young children use speech to solve a problem they do not immediately see the solution for, by verbalizing what they think they should do (Vygotsky, 1978).

Higher mental processes emerge from social processes through 'mediation', which means that a stimulus is changed by a person's response to it, using tools and signs (Vygotsky, 1978). Through internalization of concepts, children learn to use signs when communicating with others. Vygotsky mentioned the example of learning the use of pointing. Pointing to something is preceded by grasping for it. When an adult responds to this by giving the object that the child grasped for, the child will learn that grasping is interpreted as pointing and will start to use this sign deliberately for pointing (Driscoll, 1994).

Another concept that Vygotsky described, related to the emergence of higher mental processes, is that of the 'Zone of Proximal Development' (ZPD). In his experiments, Vygotsky focused on children's process of problem solving instead of on the end result. He was interested in discovering the rudimentary beginnings of new abilities, which induced his theory about the ZPD. The ZPD is 'the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers' (Vygotsky, 1978, p. 86). Vygotsky recommended a dynamic assessment of a child's skills, in which the interaction between the child and the examiner does not end when the child responds, especially if the child responds incorrectly. The examiner can provide support in completing a task and in this role serves as a teacher and a tester. In this manner, a child's ZPD can be measured (e.g., Tzuriel, 2001).

2.1.2 Scaffolding to facilitate learning

The role of the teacher or tutor in a dynamic assessment was described as scaffolding by Wood, Bruner, and Ross (1976). Scaffolding refers to 'a tutorial process; the means whereby an adult or "expert" helps somebody who is less adult or less expert' (Wood et al., 1976, p.89). The characteristics of a scaffold, as known in building constructions, are similar to those of an ideal instructor: providing support where needed, functioning as a tool and extending the range of abilities to accomplish a task otherwise not possible (Driscoll, 1994). The learner can be aided in a task to bridge the gap that exists between the actual skill level and the desired skill level. When learners have become more proficient, completing a task that they could not complete before on their own, the guidance can be withdrawn (Granott, 2005).

Wood et al. (1976) investigated their ideas on the scaffolding process in an experiment with 3-, 4- and 5- year old children. According to the authors, to be able to test the scaffolding procedure, the task should be both entertaining and challenging, it should be sufficiently complex but not lie completely beyond children's capabilities, it has to be feature rich, and it should enable children to apply the gained knowledge repetitively in the task. They used a pyramid construction task from complex, interlocking constituent blocks. Children individually completed the task, while a scaffolding process was applied by the experimenter functioning as the tutor. One of the findings was that three year old participants recognized a correct solution almost as often as the four year olds, even though they were not yet able to construct the correct solution as often as the four year olds were. The authors concluded that 'comprehension precedes production'. Thus, children are able to recognize what the solution is, before they can perform the operations necessary to reach the solution. Wood and colleagues argued that this is one of the paramount conditions to make a tutoring process effective, eventually enabling the learner to produce the correct solution without assistance (Wood et al., 1976).

Following from their study, Wood et al. (1976) listed a number of maxims with respect to the role of the tutor. For example, the tutor has to ensure the problem solver's interest in the task. Also the tutor can reduce the degrees of freedom in the task, by simplifying or reducing the required acts to reach the solution. Furthermore, the tutor may help the learner in keeping him to the pursuit of an objective, within the limits of personal interests and capacities. Also discrepancies between the learners' produced outcomes and the desired outcome can be marked by the tutor, providing information about the discrepancy between both. Frustration control was also considered to be part of the tutor's role, as well as providing demonstration (Wood et al., 1976). Although these maxims of scaffolding were developed for a human tutor, they should be useful to develop electronic tutoring.

Applying scaffolding principles in educational technologies has been subject of research for over almost two decades now (e.g., Sherin, Reiser & Edelson, 2004). Applying these principles, that were initially described for human tutoring, to the design of instructional technologies is not straightforward (Sherin et al., 2004; Granott, 2005). One of the problems is how to measure a child's ZPD. It may be difficult to interpret a child's responses appropriately (Wood, et al., 1976) and the behavior that can be measured may not provide upper and lower boundaries of a child's ZPD. Also, the ZPD is not static; it changes while the child performs a task, or a number of tasks. To match the required dynamic responses of a computerized system to the developing skills of a child is challenging. As a consequence, it is difficult to define how appropriate responses to the learner's input can be implemented in computerized scaffolding (Granott, 2005). Observable behavior of the child and interpretation thereof is different when the computer replaces the human tutor. For example, recognition of emerging skills may be difficult, based on data gathered by a computer system alone (Valsiner, 2005). In the original scaffolding procedure described by Wood et al. (1976), this was part of the role of the tutor, who can use his or her expertise to determine a child's skill level to decide when scaffolding is needed. It should be noted though that even a human tutor may not be able to determine the child's ZPD.

Recent research however, has indicated that it is possible to use electronic scaffolding for a dynamic assessment of a child's learning potential (Resing & Elliott, 2011; Resing, Steijn, Xenidou-Dervou, Stevenson, & Elliott, 2011; Resing, Xenidou-Dervou, Steijn, & Elliott, 2011; Henning, Verhaegh & Resing, 2011).

An approach to designing electronic games that are adaptive to the progress that a learner demonstrates is described in Section 2.4.4. An illustration of this approach is the development of the game 'Tap the little hedgehog', described in Chapter 5.

Requirement (1) *the content should be adaptive to the progress in the child's development,* as stated in Chapter 1, is based on the theory of the ZPD and scaffolding.

2.1.3 The role of motivation in learning

The concepts of scaffolding and the ZPD relate to another important aspect influencing the effectiveness of learning tasks: the motivation of the child (Deci, Koestner & Ryan, 1999; Lepper, Corpus & Iyengar, 2005). Persistence to complete a task, or even to start one at all, can be facilitated by making the activity attractive and motivating to the child. Therefore it is useful to consider which mechanisms may enhance motivation.

A natural drive for basic needs (e.g. food) explains behavior to some extent, though cognitive processes play an important role as well (e.g., Driscoll, 1994). Curiosity and a drive for competence towards mastering the environment cause humans to explore their environment already at a very young age (e.g., Piaget, 1952). Deciding to engage in a learning task may be caused by a person's drive for competence as well, as there is no direct basic need that can be fulfilled with such an activity (Driscoll, 1994). Once engaged in a task, again motivation is needed to complete it. Students may view successful completion of a learning task either as something that they can control, determined by their skill, or as something that is determined by chance and thus out of their control. In the latter case, students are less likely to be motivated to engage or persist in a learning task (Driscoll, 1994).

Research has focused on measures that can be taken to enhance children's motivation (e.g. Deci, Koestner & Ryan, 1999; Lepper, Corpus & Iyengar, 2005). Tangible rewards, such as candies or medals, are extrinsic rewards. Extrinsic motivation results from such external reinforcements of behavior (Gleitman, Fridlund & Reisberg, 1999). However, the use of this type of rewards was found to undermine intrinsic motivation, especially in young children (e.g., Deci, Koestner & Ryan, 1999). When the person values the activity itself, this is called intrinsic motivation (Deci, Koestner & Ryan, 2001). The exact difference between the type of rewards or processes that can be considered to be of an intrinsic or extrinsic nature is debatable. However, this debate is outside the scope of this dissertation.

From a developmental perspective, research focusing on school performances has shown that intrinsic motivation tends to decrease with age when tested from grade 3 to grade 8 (Lepper, Corpus & Iyengar, 2005). Furthermore, extrinsic motivation was found to be negatively correlated with academic outcomes. These findings demonstrate that the motivation to learn in the current school systems decreases (Lepper, Corpus & Iyengar, 2005; Malone & Lepper, 1987), compared to the motivation that children show when they are younger (e.g. Piaget, 1952).

How to make learning fun has been addressed by the classic studies by Malone, and Lepper (Malone, 1980; Malone, 1982; Malone & Lepper, 1987). Malone (1980; 1982) analysed computer games to understand which elements made such games fun, to be used for creating engaging instructional environments.

From the motivational factors that were identified, the challenge that is offered was considered to be one of the most important factors in enjoyable game play (e.g. Malone, 1980; Malone and Lepper, 1987). However, this was only found to be the case if the challenge optimally meets the skills and abilities of the player. This means that the task should not be too easy on the one hand and yet appears to be manageable on the other hand (Malone, 1980).

Malone (1980) provided a detailed set of heuristics for the design of fun educational games. The heuristics were divided in three categories: challenge, fantasy and curiosity (Malone, 1980; Malone & Lepper, 1987). Challenge relates to the following aspects, according to Malone and Lepper (1987):

- *Goal*. An activity should have a clear goal. Preferably short-term and long term goals are included. In games this can be achieved by means of multiple level goals and scorekeeping during the game.
- *Uncertain outcome.* The outcome of reaching the goal should not be certain. A variable difficulty level can be part of this, as well as successive layers of complexity, incomplete or hidden information (e.g. the game Hangman) and randomness.
- *Performance feedback.* In order to provide a continued challenge, users should receive feedback on their performance. This supports users to reformulate goals during a game. The provided feedback should be frequent, clear, constructive and encouraging.

Malone (1980) used the word fantasy as what can also be called a 'theme'. He defined two types of fantasy that can be used in a game: extrinsic and intrinsic fantasy. Extrinsic fantasy is independent of use of the skill that is trained, while intrinsic fantasies form an integral part of the skills trained. An extrinsic fantasy is generic and can be used to train different skills, such as a car race theme to train math, though the 'hangman'-theme could also be used. Intrinsic fantasies usually include real-world situations, like nowadays' serious games often do as well (i.e. virtual worlds), for example to train employees of emergency departments. The fantasy then cannot be generalized to train different skills. Transfer of the gathered knowledge to real-life situations is more likely with intrinsic fantasies, because they allow for practicing situations that may be difficult to train in real life (Malone, 1980; Malone & Lepper, 1987). In Chapter 5 of this dissertation a study is presented where an extrinsic fantasy is used to develop the game 'Tap the little hedgehog'.

Besides the training effects, Malone (1980) mentions emotional aspects of fantasies. He argues that options to personalize the fantasy based on the user's interests will provide a broader appeal to games. Thus, besides tailoring the challenge, also the theme should be tailored to the user to increase its appeal and educational effectiveness.

Like the challenge and fantasy game aspects, also curiosity requires some kind of balance. The game environment should provide enough surprises to stimulate curiosity, but it should also be comprehensible for the player (Malone, 1980). Malone distinguished sensory curiosity and cognitive curiosity. The first may be enhanced by visual and audio effects, while the second depends on 'a desire to bring better "form" to one's knowledge structures' (Malone, 1980, p. 166). To support cognitive curiosity, feedback should be surprising to enhance engagement, but also constructive to enhance learning.

Another more recently proposed model described motivational aspects in terms of human states, instead of game design heuristics. As the outcome of a thought-exercise on the mechanics of fun in playful learning, Fontijn and Hoonhout (2007) proposed a model on the 'core sources of fun'. They argued that learning can be seen as 'a second order survival activity', as it does not meet immediate survival needs (i.e. food), but it does help to become skilled at activities (i.e. acquiring food) that are necessary to meet the first order needs. Fun was suggested to be the evolutionary mechanism to reward increasing the ability to survive, i.e. learning, and make it intrinsically motivating. Hence Fontijn and Hoonhout (2007) proposed three targets for learning, based on second order survival needs: developing skills, gaining knowledge and bonding. Mapping these onto games and toys resulted in three core sources of fun: accomplishment, discovery, and bonding. Accomplishment has to do with perceived progress towards goals and meeting them (e.g., Malone & Lepper, 1987). Discovery relates to a child's curiosity to explore the world around him (e.g., Piaget, 1952). Bonding refers to an individual's sense of recognition and affirmation by a group and a sense of belonging, 'the need to be needed' (Fontijn & Hoonhout, 2007).The authors argue that the best toys and games combine at least two of these core sources of fun, and that these aspects should be considered as independent of each other. Furthermore, 'enhancement' factors were distinguished that are assumed to enhance the effectiveness of the core sources of fun, but on their own they do not provide fun. There are at least two and possibly three enhancement factors according to the authors: fantasy, aesthetics (of the interface, but also in elegance of game rules) and possibly physicality. Physicality relates to an interface that is physically engaging. However, it is unclear why this enhances fun. The reason may be related to relevance, multimodality or simply that the added complexity enriches the challenge (Fontijn & Hoonhout, 2007).

The definition that Malone and Lepper (1987) used for toys and games is very similar to the definition of intrinsic motivation: games and toys are used 'for their own sake with no external goal' (Malone & Lepper, 1987, p. 234). This indicates their usefulness as engaging learning environments. Computer-based serious games offer a learning environment that allows for different ways of supporting children's motivation, including the aspects described above. Although these games are currently used in a broad range of learning and training domains, including the military, health care, city planning, engineering, politics and emergency management, such as fire departments, they would also fit very well in schools (Ulicsak & Wright, 2010). They offer a way to learn in the real world in an enjoyable manner and can be fully tailored to individual learning needs by taking into account a child's ZPD.

Requirement (2) stated in Chapter 1, that *content should be self-motivating* is based on the theories on motivation described above.

2.2 Tangible User Interfaces: tangible and embodied interaction

2.2.1 Definition

Physicality or tangibility as a mode of interaction with computers, has received increased research interest over the past two decades. The definition of Tangible User Interfaces (TUIs) has been, and continues to be, a topic of discussion (e.g., Fishkin, 2004; Shaer & Hornecker, 2010). Traditionally, tangible user interfaces are said to 'couple physical representations with digital representations, yielding user interfaces that are computationally mediated but generally not identifiable as "computers" per se.' (Ullmer & Ishii, 2000, p.917). TUIs integrate input and output in physical objects that represent digital information themselves (Ullmer & Ishii, 2000). Whereas graphical user interfaces (GUIs), such as a PC mouse and a screen, separate input and output modalities, TUIs seamlessly integrate control and representation. Furthermore, TUIs allow for using a wider range of human skills and abilities such as perception, motor skills and emotion (Dourish, 2004).

A 2D taxonomy has been proposed by Fishkin (2004) with the axes 'embodiment' and 'metaphor' as a way to classify TUIs. Embodiment was described as the closeness of input and output in an object, ranging from full to distant (Fishkin, 2004). As an example of full embodiment, clay was mentioned: the input is the output. Tabletop applications are positioned as nearby on the embodiment axis, as the input typically takes place nearby the output. On the metaphor axis, ranging from none to full, none means that there is no relation between the action with the TUI and the real world. As an extreme example computer command lines are mentioned. Full means that no metaphor or analogy is necessary, the virtual system is the physical system, so manipulating the first, means the same happens to the second (Fishkin, 2004). Note that this taxonomy parallels Malone's description of Intrinsic Fantasy (Malone, 1980, 1987): examples of full embodiment and full metaphor are similar to a fantasy theme that requires no transfer to the real world, as it is already so close or even identical to it.

2.2.2 TUIs for education

As described before, children's explorative play with physical objects supports learning. Piaget noted that the first stages of cognitive development are strongly influenced by children's interactions with physical, tangible objects (e.g., Piaget 1936/1977; Goswami, 2008). Vygotsky (1978) stressed the importance of activity and the use of tools in child development. Bruner (1973) called for teaching that takes into account the children's natural thought processes by giving them the opportunity to progress beyond their, in his words, primitive modes of thinking through confrontation with concrete materials. These concrete activities, it was argued, would eventually convert into mental representations (Bruner, 1973, p. 414). Van Parreren and Carpay (1980) and Reid (2001) have also stressed the role of tangible objects in education and teaching. Other researchers stated that interaction with tangible objects will benefit a child's learning in general (Goswami, 2004; O'Malley & Stanton-Fraser, 2004; Khandelwal, 2007) because manipulation of objects can support the child in constructing mental representations of the world around him or her, and in creating knowledge about physical events.

Scholars like Dienes (1964), Montessori (1912) and Fröbel (1826) have advocated the use of tangible materials in schools (Manches, 2009). They developed sets of materials for classroom use, such as Dienes' multi-base arithmetic blocks, intended to facilitate comprehension of elementary mathematics by the formation of 'qualitative structures', for example the concept of number (e.g., Piaget, 1976). According to Montessori, playing with physical objects supported self-directed, purposeful activities (O'Malley & Stanton-Fraser, 2004). The designed materials were quite simple, such as wooden cubes and colored rods to teach different topics. These materials are still used in classrooms today, in the early school years, but also for older children (Freer Weiss, 2006; Moyer, 2001).

A student of Piaget, Papert, launched the paradigm of educational computing which was referred to as 'constructionism' (Papert, 1980). Papert created a turtle robot that could be controlled by Logo programming code. The activity of programming the code for the turtle would enable children to learn geometric concepts more easily than previous more abstract math teaching methods, and support children in making their implicit knowledge explicit. Building on Papert's work, Resnick introduced the concept of 'digital manipulatives': 'computationally-enhanced versions of traditional children's toys' (Resnick, Martin, Berg, Borovoy, Colella, Kramer & Silverman, 1998, p. 281). The work of Papert and Resnick was part of the development of programmable construction kits for children, i.e. Lego Mindstorms®, designed to enable children to learn how to program.

The use of TUIs to teach different educational topics has been explored in a broad range of studies. The typical learning domains associated with TUIs include developing language skills, training computer programming, teaching about molecular biology or chemistry, and also about dynamic systems such as factory processes (Marshall, 2007).

Examples of exploratory projects include the Teaching Table, an interactive tabletop device for prekindergarten children, developed to train math-related skills (Khandelwal & Mazalek, 2007). Several tangible storytelling applications to train verbal skills and develop vocabulary have been designed, such as StoryToy by Fontijn and Mendels (2005) and Storytent by O'Malley and Stanton-Fraser (2004). Teaching literacy skills to children with multiple disabilities was investigated by Hengeveld (2011) with LinguaBytes, a tangible learning system. Supporting children's collaborative play (Africano, Berg, Lindbergh, Lundholm, Nilbrink & Persson, 2004) and physical activity by outdoor games for children using tangible devices are other examples of TUI-based research (e.g. Soute, Markopoulos & Magielse, 2009; Verhaegh, Soute, Kessels & Markopoulos, 2006).

These studies all aimed to use TUIs to enhance children's learning or playing process. Based on theory and such exploratory studies, several possible benefits of TUIs for learning have been described. For example, TUIs are assumed to support playful learning, which enhances children's engagement in a learning task. Furthermore, they are claimed to be a more accessible and direct interface than for example PC-based learning applications, and to support multisensory learning as well as collaborative play (Manches, 2009; Marshall, 2007). Also, it has often been suggested that the use of tangible task formats in education has clear advantages over PC-based tasks (Antle, 2007; Khandelwal & Mazalek, 2007; O'Malley & Stanton Fraser, 2004; Verhaegh, Fontijn, & Hoonhout, 2007).

Requirement (3a) that *the emphasis lies on training the intended skills*, stated in Chapter 1, builds on the accessibility claim with regard to TUIs. The use of TUIs can avoid unwanted cognitive load by providing direct access to a task. This will be explained in more detail in Chapter 3.

There is, however, limited empirical evidence to substantiate these claims with regard to their exact effects on performance in learning tasks (Marshall, 2007; Olkun 2003; Olkun, Altun, & Smith, 2005).

2.3 Introduction of TagTiles

2.3.1 What is TagTiles?

TagTiles is a tabletop TUI, a console (see Fig. 2.1) that was developed by Serious Toys BV (www.serioustoys.com) to support independent learning. Children interact with TagTiles through tangible objects. The TagTiles board supports many different types of educational games addressing language, math and cognitive skills. To start using the system, children can log in by means of a name tag that is identified by the board, to start immediately with one of the games, or to continue where they left off last time.

The console includes a tabletop sensing board with an array of LED lights underneath and audio output. The size of the board is A3, in cm this is 30(W) x 40(L) x 5.5(H), with a detection area of $24cm^2$. It is made out of hard plastic material, which is robust and suited for classroom use. The playing objects include RFID tags to enable their localization and identification on the board. Any object can be used, depending on the requirements for the game, such as miniature boats, cars and puppets, and also more abstract shapes such as cubes and circles.

Figure 2.1. The TagTiles console with some playing objects.

Removable plastic foils are used on top of the TagTiles board to display different game layouts. Various foils and objects can be used to play different games. There is a slot at the top of the board where assignment cards can be placed upright. The games are run from an SD card which is inserted in a slot at the backside of the console. The child receives instructions on how to play the games via sound. During play, the system provides feedback on the child's actions on the board. This feedback can be auditory or visual, the latter through the LED lights in the board, or a combination of both.

Manipulations of playing objects on the board can be logged by the system in a detailed manner. The measurement is automatic and objective, making it consistent across children, making comparing between children easier. Because the measurement takes place during play the latter can be meticulously followed, as the logs provide precise information on the location and the time of objects being place on the board. Based on this information, it can be identified afterwards how a task was solved. As such, the logs are a rich source of information that allows for interpretation on different levels, which can be useful for different purposes and users.

For example, teachers may want to receive a report of the activities that a child has completed on the board. A remedial teacher may require a more detailed report of a child's performances, to see how a child performed on different parts of a task. This can provide more precise indications of where a child needs extra training. A therapist may wish to have a report on how often a task was performed, in case the child trained without the presence of the therapist. For research purposes the recorded information is very useful as well, as it enables relatively easy evaluations of new applications, requiring no additional recordings such as video during a test.

Recorded performance information is also used by the system itself. Completed tasks or levels, and also speed of play can be taken into account in setting the difficulty level of an activity. This allows for offering content that fits with a child's ZPD, and to apply scaffolding when needed. How this may be achieved, is explained in paragraph 2.4.4.

2.3.2 Educational applications of TagTiles

The features of the console enable many types of learning, for example, rote learning (e.g. of math tables), training language related skills such as understanding letter-sound correspondences, but also facilitating understanding of concepts, such as spatial relations between objects. Different games have been created to investigate whether the console can be used to address skills in the areas of cognitive, motor and social development. In the following sections three game applications targeting the three developmental domains are described.

Cognitive skills: original TagTiles

The game TagTiles (Verhaegh et al., 2007) was initially developed as a tool to investigate the factors that provide fun in a tangible educational game. It should be noted that the name 'TagTiles' was initially used for both the console itself and the game. The game, implemented on an early TagTiles prototype (Figure 2.2), was designed to train mostly cognitive skills, such as spatial insight and memory.

Two players can participate in the game in which they both need to copy patterns that are displayed on a central LED array of 8x8 fields. Copying is done by tagging the corresponding fields on a sensing board in front of each player (see Figure 2.2). The game is played in competition and the one who first copies the pattern accurately, gains a point.

Figure 2.2. Children playing the original TagTiles game. A pattern that needs to be copied is displayed on the middle grid. With the colored cubes the players tag the fields to copy the pattern. Picture of TagTiles demonstration at 'TU/e publieksdag': open house event at the Eindhoven University of Technology.

An evaluation of the game (Verhaegh et al., 2007) with thirty-eight 10-12 year-olds indicated that TagTiles is a fun game to play and that the game can easily be adapted to the skill level of the players by changing the patterns offered or the task connected to the pattern. For example, memory skills are addressed if the players need to reproduce the pattern after it is removed. The evaluation results also provided clues on how to create different challenges for different players within the same game. In addition, it was observed that playing the game together provided social benefits. Children seemed to learn from each other's strategies in completing the task and they would help one another if one was lagging behind (Verhaegh et al., 2007).

Motor skills: TagTiles for therapy

The test with the original TagTiles game showed the potential value for training cognitive skills. Moreover, it was observed that motor skills could be supported as well by tailoring the game objects to elicit specific movements. This can be extremely useful in for instance therapeutic settings (Li, Fontijn & Markopoulos, 2008).

Cerebral Palsy (CP) refers to a number of neurological disorders that appear in infancy or early childhood and affect body movements and coordination permanently. Although CP cannot be cured, therapy can improve the child's abilities and self-confidence (Ekström Ahl, Johansson, Granat & Borgren Carlberg, 2005).

In a participatory design process involving children with CP and their occupational therapists at a number of clinics, information was gathered on the current practice of therapy and problems that occur. Current therapy was found to be tailored to the specific needs of each child, using all kinds of materials to elicit desired movements. For example, to train a pincer grasp, children are asked to move paper-clips from one side of a card to the other. Issues noted with the current therapy practice included compensating for difficult movements, avoiding the use of the affected hand or arm instead of training it. Also, a lack of cognitive challenge sometimes caused the therapy to be boring, risking a loss of the child's motivation. In addition, due to the mostly individual setting, current therapy misses out on the potential benefits of a social setting.

A number of games were developed to support the therapy of children with CP (see Li et al., 2008, for a description) to train specific movements that these children need to train, such as elbow extension (see Figure 2.3).

The evaluation of the game prototypes in the clinics showed that the children (seven participated, aged 4-11 years) spontaneously made a number of the desired movements while playing, such as extension of the elbow. The children said to enjoy playing the games and this was also reflected in the obvious effort they put into playing them. Children indicated to prefer games that provided a cognitive task that was challenging for them (Li et al., 2008).

Figure 2.3. TagTiles game for therapy. Spontaneous extension of the elbow of the affected arm motivated by the game.

Social skills: Playground Architect

To investigate the potential of the console in the domain of social skills, another design project was initiated. Several interviews with school teachers provided insight into the issues that children have in the social skill domain. The interview results showed that many children could benefit from becoming more assertive and less shy, and it was decided to create a game that can support this. Several game concepts were developed and eventually 'Playground Architect' was selected (Hendrix, van Herk, Verhaegh & Markopoulos, 2009).

In Playground Architect (see Figure 2.4), three to five children can participate. One of them takes the role of Architect. The other players are Builders. The Architect's role is specifically intended for a shy child. He or she receives the Architect's pawn and the Builders have the playground objects which they have to place onto the board according to the Architect's instructions.

The Architect is the only one who can access a set of narrative instructions that describe the client's wishes. These instructions are played back via the Architect's headphones. The instructions involve choices that are to be made by the Architect or by all players together, depending on the Architect's preference. If the Architect makes the decision alone, this can be seen as a sign of assertiveness or self confidence.

Forty children (mean age 9.5 years) participated in an evaluation of the game that took place at their school. Before the evaluation, they were all tested for Dominance/Shyness via a teacher questionnaire, based on which the shyest children were placed in the Architect's role.

Figure 2.4. Children playing the Playground Architect game. The board lights up to show that all objects have been placed correctly so far.

An analysis of speech during the recorded play sessions showed that the shy children (the Architects) talked at least as much as the less shy children (the Builders). Peer acceptance was also measured and in many cases it increased already after only a single round of play. Reviewing the play sessions with the children's teachers gave overall very positive reactions (see Hendrix et al., (2009) for complete results).

2.3.3 Integral skill development

The three studies indicate that one skill set can be used in optimizing the development of another. For example, the original TagTiles game, aimed at cognitive skills, employs motor skills (placing the blocks) and a social component (competition) to increase motivation. Similarly, the CP games, aimed at fine motor skills, use a cognitive challenge to tune the overall challenge of the game. The Playground Architect game, aimed at social skills, also uses a tunable cognitive challenge. Based on the evaluations, it is expected that fine motor skills will be improved by the TagTiles game aimed at cognitive skills. It can also be expected that the spatial skills of children improve with the CP games aimed at fine motor skills. Furthermore, collaborative games like Playground Architect can easily be augmented to add linguistic or math challenges. This suggests that the TagTiles console is useful for an integral approach towards skill development.

The three studies also indicate that the TagTiles console may be used in conjunction for assessment, learning and therapy. One and the same game can be used to assess the skill level of a child and tune the challenge based on that assessment. If the child is on par with its peers (or above) this would constitute regular (or accelerated) learning, if the child is below par this would deliver a remedial effect. Key is that the game can be tuned to the specific needs and abilities of each individual player, and this can only be done based on domain expertise. Because the exercises can be presented in the form of attractive games, the children are intrinsically motivated to use them. Finally, such tools can be used by the children unsupervised and hence as easily at home as in a more formal setting. Currently, we have not found any other tools than tangible electronics with the ability to deliver this combined set of benefits in an integral manner.

Requirement (3b) *the content development process should involve domain experts*, as defined in Chapter 1, is based on the experience gathered in game development studies for the TagTiles console. These studies illustrated that the involvement of domain experts is required in the process of creating effective educational content.

2.4 Introduction to the experiments

The results of the three studies with the TagTiles games addressing cognitive, motor and social skill-domain served as a starting point for the game development experiments described in this dissertation (in Chapters 4 and 5). The focus of research was narrowed to the assessment and development of cognitive skills. The target group for this work and the applied experimental measures are described in the current section. Also the applied approach to designing a game that is adaptive to a child's performance is described.

2.4.1 Target user group

Children attending regular primary schools, aged 7-12 years were included in the studies. For one, as the use of tangible learning materials is common in primary schools, especially in the early years, TUIs fit naturally in the curriculum.

Second, the age of 7 years was taken as a lower boundary for developmental reasons. It is known that children below the age of 7 years often show instable performances in (intelligence) tests and we anticipated repeated testing, comparing subsequent test results. Also, motor skills were found to have reached a mature stage, at least of the most fundamental movement skills, at about the age of 6 years (Gallahue & Ozmun, 2002).

Further, language skills are well-developed, which justifies the use of interviews and questionnaires, albeit using questions that do not include complex words as children's vocabulary is still limited at this age.

Processing speed determines performance on many cognitive tasks. Research has shown that speed of processing is still developing in the target age group, and full processing maturity is not reached until approximately the age of 19 years (Sternberg & Berg, 1992).

Participants from a particular school level were included for each test, ensuring a limited variety between the children's intellectual levels. The only other exclusion criterion was color blindness, which was tested beforehand.

2.4.2 Experimental measures: intelligence tests

WISC-III^{NL}

The Wechsler Intelligence Scale for Children (WISC) is a commonly used test to measure intelligence. It is also used to diagnose learning disabilities, usually as a part of a larger test battery.

A selection of subtests of the Dutch version of the Wechsler Intelligence Scale for Children, third edition NL (Wechsler, 2005) was used in the study in Chapter 4. The following nonverbal subtests were included: Coding, Symbol search, Picture completion, Block design, Object assembly and Digit span. Vocabulary was the only verbal test that was included.

RAKIT

The RAKIT (Revisie Amsterdamse Kinder Intelligentie Test; Bleichrodt, Drenth, Zaal, & Resing, 1987) is an intelligence test also consisting of several subtests assessing verbal and non-verbal intelligence. Part of the Memory Span subtest to assess sequential visual memory for concrete figures, and the Hidden Figures subtest was applied to assess visual search. In this test, children have to find a hidden figure in a complex drawing.

Raven
Raven's Standard Progressive Matrices (Raven, Raven & Court, 2004) is a commonly applied non-verbal test and it appeals on deductive reasoning with patterns, to assess IQ from a pattern completion and reasoning perspective. The test has a pencil-and -paper format.

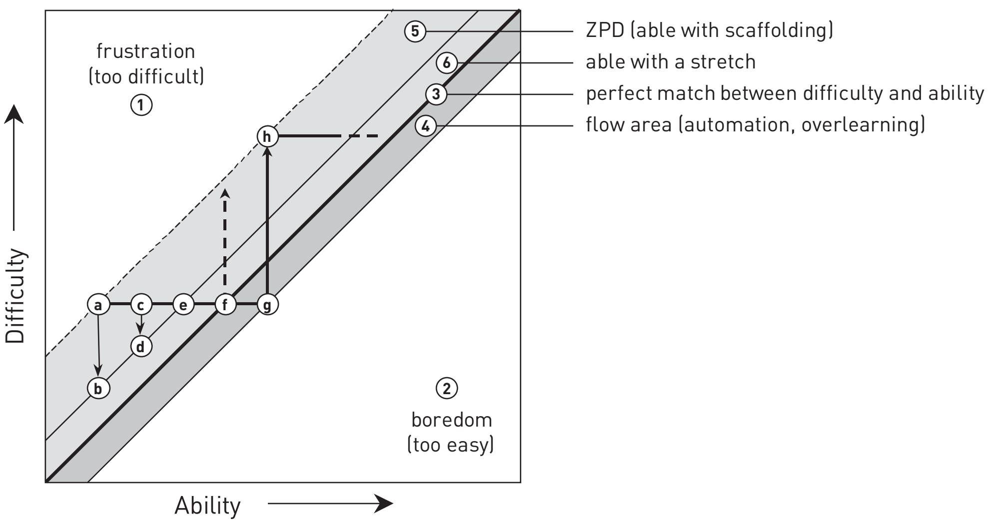
2.4.3 Experimental designs

The studies in Chapters 3, 4 and 5 each describe a field test at schools using so-called within-subjects experimental designs. This means that same children were placed in different conditions in each study, enabling comparison of the results. Having the same children in different conditions reduces the error variance associated with individual differences, and increases the power of statistical tests. In Chapter 5 however, children were tested several times in one condition to assess the development of game performance of children over three sessions.

2.4.4 An approach to designing adaptive educational games

The game 'Tap the little hedgehog' is described in Chapter 5. In the current section some background is provided to the adaptation of the game to the child's progress. The theory of the ZPD combined with the concept of scaffolding is deemed very useful as a theoretical basis for the development of electronic educational applications. First, if the tasks offered to the child are within the ZPD, new skill development can be facilitated (Vygotsky, 1978). An important risk is perhaps the reverse: (far) outside the ZPD learning might not take place. For example, if the task is far beyond the child's current abilities, even tutoring may not make a difference because the task is far too difficult for the child to grasp. In case the task is too easy, the child might stop because of boredom. If a task is far below the ZPD it is expected that at most automation of skills will occur. Second, offering tasks that are within the ZPD may make it more likely that the child experiences the task as fun to do, because frustration and boredom can be avoided. This is related to the concept of 'Flow'; a concept that was coined by Csikszentmihalyi (1990). Flow was described by him as 'being completely involved in an activity for its own sake [...]. Your whole being is involved, and you're using your skills to the utmost' (Geirland, 1996).

How the ZPD and scaffolding in an educational task can be used, is illustrated in Figure 2.5. This approach (Fontijn & Verhaegh, 2008) is intended for educational game design, and it is based on practical insights gained from educational game design experience. The depiction is similar to earlier representations of the ZPD and also resembles representations of the Flow theory (Csikszentmihalyi, 1990). More details are provided in this figure for its practical use in designing educational tasks. The ideas represented can guide the implementation of scaffolding and level adaptivity into an electronic educational game. It should be noted though that they were not formally tested in this dissertation.



ZPD related areas

(1) The level of difficulty is too high for the child to master or grasp even with help.
(2) The level of difficulty is too low for the child to be interested. The task would just be work.
(3) Perfect match between level of difficulty and child's ability; there is a proper challenge and the child can complete the task without help.
(4) The 'flow' area. The child is in an area where it will optimize its performance on completing the task.
(5) The ZPD area. The child needs help to complete the task.
(6) Area where the child can solve the task on its own but by stretching, e.g. by trial and error.

Application of scaffolding in the learning process

(a) The challenge is too large for the child; substantial help is required to bring (b) in reach of the child.
(b) This is the level at which the child can actually solve the task but it will have been helped to complete these tasks, as they are more advanced than what the child would have achieved on its own.
(c) The task is still too difficult but the child has learned, and less help is required.

(d) The child is actually solving the same task at a higher difficulty level, i.e., higher than at level b, but with help.
(e) Still the same difficulty level but now the child has learned enough to complete the task on its own.
(f) The actual ability of the child matches the difficulty level of the task, the child knows how to solve the task; at this point one may consider offering the next difficulty level.
(g) The child is comfortable with the task; we can observe it can solve the task flawlessly but with increased speed, this is definitely the moment to increase the difficulty level or the child will lose interest.
(h) Next stage, scaffolding is again required.

Figure 2.5. Matching the difficulty of a task with a child's current abilities. The figure illustrates how the child progresses through the ZPD and related areas while performing a task (areas 1-6), and at which points scaffolding and/or tasks of a higher or lower level of difficulty should be offered (points a-h).

The figure shows a number of principles. First, it illustrates at which moments in the child's progress through a task it makes sense to provide some kind of support. This would only be in the ZPD, area (5). Points (a) – (g) explain how scaffolding can be applied in this area. With scaffolding the actual level of difficulty of the task is reduced and brought within reach of the child. Second, it shows when the difficulty level should be increased or decreased, ensuring continued interest and avoiding boredom or frustration of the child. Preventing boredom requires providing a higher difficulty level at least before reaching point (g), while frustration can be avoided by applying scaffolding at least before reaching point (h). Finally, the figure illustrates the need for precise and continuous measurement of the child's performance, in order to enable implementation of the first two principles. This is why automated electronic scaffolding may be preferred over human scaffolding, especially when it concerns a task that requires fine grained, though machine-measurable steps that may be detected more accurately by computer monitoring.

Going from area (3) through (4) to (2), we assume that the speed of learning decreases. One could argue that in area (4) the focus is not on learning new things but on learning to perform a task better. In area (6) a child does learn new things. New strategies or abilities are needed to solve the task. The distance between actual and required ability is such that the child can bridge the gap on its own. In area (5) the task becomes too difficult for the child to come up with such strategies but it will still be able to complete the task with help (scaffolding). In area (1) the task is too difficult to complete with the current level of performance that the child demonstrates.

It first has to learn new skills to be able to reach that level, even if help is provided. It is assumed that the child learns fastest in area (5), the ZPD, though provided that the scaffolding is applied in the right manner (Fontijn & Verhaegh, 2008; Henning, Verhaegh & Resing, 2011).

3

Study I

Playing with blocks or with the computer? Solving complex visual-spatial reasoning tasks: Comparing children's performance on tangible and virtual puzzles

Verhaegh, J., Resing, W.C.M., Jacobs, A.P.A. & Fontijn, W.F.J. (2009). Playing with blocks or with the computer? Solving complex visual-spatial reasoning tasks: Comparing children's performance on tangible and virtual puzzles. *Educational & Child Psychology, Vol. 26* (3), 18-29.

Abstract

This study examined the solving of a visual-spatial reasoning puzzle task by children. The goal of the experiment was to measure how the physical form in which an educational task was presented influenced children's task performance. All children were administered two sets of randomly chosen puzzle tasks. The order in which the task format was presented differed: children in condition one were presented tangible puzzle tasks first; in the second condition tasks in a virtual format were presented first. We compared the task performances of 26 children aged 5-7 in both conditions. The results show that, independent of the order of the task format, task completion times for the tangible version were significantly shorter than those for the virtual version. In addition, we found that children engaged in far more overt problem solving behaviour when using the tangible version; this most likely helped them to find the correct solution quicker. These findings support our hypothesis that tangible interfaces are more appropriate for training visual-spatial reasoning in this age group than are their screen-based virtual counterparts. The latter may require additional, underlying skills to operate the interface. Indeed, the results suggest that, for the virtual format, the variations in the task completion times between task types can be attributed to the number of puzzle piece rotations required.

3.1 Introduction

Complex problem solving in children is often measured by administering puzzle tasks such as Block Design, the Tower of Hanoi, or verbal problem stories that have to be solved logically, e.g. by induction. In this contribution, we focused on the solving by children of visual-spatial tasks that can be used in an educational setting to measure and enhance complex problem solving. To be able to develop effective educational tasks, knowledge about the influence of task components is essential. One important task component is its interface, i.e. the physical form in which the task is presented to the child. To study the effect of different task formats on complex problem solving in children, we set up a study in which two versions of a puzzle task were compared: one with tangible materials, which we called the tangible version and one with a computer interface, the virtual task version. We also examined if children's problem solving performance changed over the succession of tasks.

As the world is becoming increasingly computerized, many computer-supported solutions for the educational domain are developed. Assessment and instructional settings for children have been computerized (e.g., Tzuriel & Shamir, 2002). Until recently, task design for educational assessment focused mainly on screen-based applications, but now we are witnessing explorations of computerized tangible tasks in combination with sensor technology. These offer interesting new options, such as allowing more flexibility in problem solving and presenting a more natural environment for assessment (e.g., Fontijn & Mendels, 2005; Khandelwal, 2006; Revelle, Zuckerman, Druin, & Bolas, 2005; Terrenghi, Kranz, Holleis, & Schmidt, 2006; Verhaegh et al., 2007).

Electronic tangibles are digitally enhanced physical objects, for example wooden blocks, in which the influence and visibility of the computer is reduced as much as possible (O'Malley & Stanton Fraser, 2004). Computerized tangibles combine the benefits that physical objects offer for learning with the benefits that computers bring. As computerized tangibles potentially enable a relatively new way of assessment and training for children, it is necessary to examine how they can be used effectively in an educational context. An important issue to study first is how the task format may influence children's task performance. A precise understanding of the influence of task components is a prerequisite for offering appropriate instruction or training to a child. Although the format in which a task is presented is only one component of a task, it may contain a relatively large part of the challenge that the child is presented with, especially if he or she finds it difficult to handle. Often the actual task to perform lies 'behind' the interface and the additional challenge this causes is unwanted. Educational tasks become more effective when the targeted skills can be directly addressed.

Tangible learning aids, for example, building blocks, have had their place in education for a long time. In many early years settings children play and build with construction blocks, and in the first years of primary education children regularly use physical objects when they have to learn to count or to measure. Tangibles, sometimes called manipulatives, are used in mathematics for rather older children also (Freer Weiss, 2006; Moyer, 2001).

Piaget noted that the first stages of cognitive development are strongly influenced by children's interactions with physical, tangible objects (e.g., Piaget 1936/1977). Vygotsky (1978) stressed the importance of activity and the use of tools in child development. Bruner called for teaching that takes into account the children's natural thought processes by giving them the opportunity to progress beyond their "primitive modes of thinking" through confrontation with concrete materials. These concrete activities, it was argued, would eventually convert into mental representations (Bruner, 1973, p. 414). Van Parreren and Carpay (1980) and Reid (2001) have also stressed the role of tangible objects in education and teaching. Other researchers stated that interaction with tangible objects will benefit a child's learning in general

(Goswami, 2004; O'Malley & Stanton-Fraser, 2004; Khandelwal, 2006) because manipulation of objects can support the child in constructing mental representations of the world around him/her, and in creating knowledge about physical events.

Tangibles are seen as providing more visual-spatial freedom than the personal computer, because dimensionality is not reduced to a flat screen. In addition, they provide children with very different tactile experiences than computer interfaces (Olkun, 2003). At a psychological level, it can be said that tangible task formats are easier to use and support learning better, due to the less symbolic representational form in which objects are presented (Antle, 2007; Price & Rogers, 2004; Ullmer & Ishii, 2001). Furthermore, by moving an object by hand, both the hand and the object move in parallel towards the place required, while with the virtual equivalent the control point of the object and the object representation are separated in space, unless a touch screen is used. The objects themselves also differ: a tangible is a concrete representation, while an object on a screen consists of a symbolic 2D-representation. In other words, with a tangible interface the task and the interface are integrated, while with a virtual interface they are separated. Hence, the latter requires additional interpretation or translation steps, just to cope with the interface.

Although the advantages of tangible materials can be defended rationally, research on the effectiveness of the use of tangible materials has not always led to results that confirm expectations. Some older research has shown that on specific tasks, such as mass or volume conservation (e.g., Bruner & Olver, 1963; Piaget, 1954, 1970), children were successful with concrete physical objects before they were able to solve the tasks symbolically (e.g., Sowell, 1989; O'Malley & Stanton Fraser, 2004). Others however, reported equal improvements in performance after training in solving tangible geometry and tangram problems versus equivalent virtual tasks (Olkun, 2003; Olkun et al., 2005).

It has often been suggested that the use of tangible task formats in education has clear benefits above non-tangible ones, such as tasks on a computer (Antle, 2007; Khandelwal, 2006; O'Malley & Stanton Fraser, 2004; Verhaegh, Fontijn, & Hoonhout, 2007). There is, however, little empirical evidence to substantiate this (Marshall, 2007; Olkun 2003, Olkun et al., 2005). Very little empirical evidence is available to support the claim that tangible materials enhance performance on complex visual spatial reasoning tasks. Earlier research, starting with Piaget's work, has emphasized the importance of the use of tangibles in children's development. However, the exact effects on learning, i.e. how the characteristics of tangibles result in development, are still unclear. Over the past decades, as new technologies emerged, new interfaces were introduced into education, e.g. the PC has found its way into the classroom.

In our broader study of the effect of tangibles on the learning process we sought to determine whether a task presented on a PC and involving mouse use, requires different skills than the same task presented in a physical format. We did this by looking at the differences between tangible and virtual task formats.

The main purpose of our study was to determine the influence of the task format on children's performances on a complex visual-spatial problem solving task. We investigated whether a tangible or a virtual format was more difficult to solve for 5-7 year old children. This age group was chosen for our study as such children are in a transitional phase in which they are getting used to working with pen and paper and are spending a decreasing amount of time working with tangible objects. In addition, many children of this age are not yet skilled PC users. Therefore we expected that controlling a mouse and understanding how to drag and drop would still be a challenge to them. When developing educational applications, it is important to take into account how a PC interface may change the educational effect of a task, especially for this age group.

In our study we took into account the effect of task format order on children's performance, and investigated possible gender effects. We also studied whether one of the two task formats led to more or different overt forms of problem solving. We compared the number of measuring activities and fitting attempts children made when completing tasks in both formats. Furthermore we examined children's performance over the sequence of tasks, to answer the question whether practice led to faster task completion (e.g., Olkun, 2003). We also investigated which version of the tasks children preferred and which version was found to be most difficult, as this may have influenced children's motivation to perform optimally (e.g., Malone & Lepper, 1987; Moyer, 2001).

3.2 Method

3.2.1 Participants

Twenty-six children (mean age 6.6 years), attending second and third grade of a regular primary school, half boys and half girls, participated in the study. The children came from one school in the southern part of the Netherlands. All children had at least some prior experience of using a personal computer. Written informed consent was obtained from parents prior to participation.

3.2.2 Materials

For the experiment we used an existing educational puzzle-set: 'Passen & Meten 1' (Fitting & Measuring 1) developed by Jegro. The puzzle-set consisted of different blocks of various lengths and colours that had to be arranged according to a coloured pattern. The puzzle-set had five puzzle types and was constructed to measure complex visual-spatial reasoning. To solve the tasks, children had to measure, rotate and compare objects of different lengths and colours, reason by exploration, and think logically. Two versions of the Fitting & Measuring tasks were used (see Figure 3.1).

The tangible version of the puzzle task consisted of wooden blocks and plastic assignment cards. The blocks had five different colours (red, yellow, blue, orange and green) and for each colour there were five rectangular blocks in sizes ranging from 1x1x1 cm to 5x1x1 cm. Children were asked to match a pattern presented on the assignment card using the blocks available. The blocks were presented unsorted in a wooden box, and were continuously visible throughout the task. To place a block, it had to be picked up and directly, or after rotation, placed on the correct spot on the assignment card.

Figure 3.1. Tangible version (left) and virtual version (right) of the puzzle task.

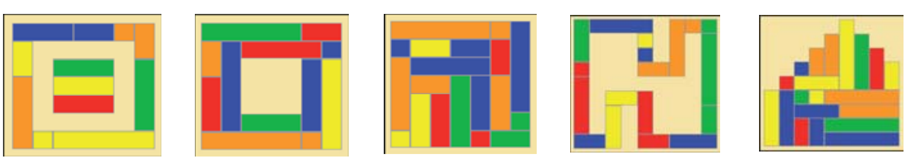
In the virtual version of the puzzle task, the puzzle was shown on a computer screen divided in two parts: on the left side of the screen virtual blocks were presented, and on the right side was displayed the pattern that needed to be filled. The virtual blocks had to be selected and manipulated with the computer mouse. Blocks of only one colour were visible at a time, already sorted by size (see Figure 3.1). Pushing a virtual button positioned underneath the blocks changed their colour. Clicking a block on the left side of the screen changed its orientation from horizontal to vertical or vice versa. To place a block on the right spot in the puzzle, it had to be given the correct orientation, dragged to the correct spot on the pattern and then dropped.

The goal for the virtual and tangible task versions was the same: to match a pattern using the blocks available. The only difference between the two versions was the type of user interface. For both versions the assignments consisted of the same five types of puzzles. Table 3.1 gives an overview of the average number of elements, colours, and horizontal/vertical transformations a child had to consider in solving each of the 5 task types, as well as examples of the 5 puzzle tasks types patterns. In Table 3.2 the main differences between the tangible and virtual task versions are outlined.

Table 3.1. The average number of elements per virtual task type including orientation and number of colours (upper part); tangible tasks 1-5 (lower part).

Task	Number of blocks	Horizontal*	Vertical	Colours
1	11.3	7.0	4.2	5
2	13.2	7.5	5.7	5
3	17.0	12.0	5.0	5
4	16.9	8.8	8.2	5
5	20.0	8.6	11.3	5

* squares are counted as horizontal as they do not require rotation



1 2 3 4 5

Table 3.2. Differences between tangible and virtual versions.

Task elements or action	Tangible	Virtual
Presentation blocks	Unsorted All blocks are visible at all times	Sorted by colour and size Only blocks of one colour visible at one time
Placing blocks	By hand	By mouse, drag-and-drop required
Rotating blocks	By hand	By mouse, clicking on block before moving

3.2.3 Design and Procedure

Each participating child was brought individually from the classroom to a separate, quiet room. The child was always asked for permission to record the experiment on video. The experiment started by explaining the test procedure.

All children were administered two sets of randomly chosen puzzle tasks. The order of task format differed: the children in condition one were presented with the tangible puzzle tasks first; in condition two the virtual versions were presented first.

Each child had to complete four puzzles per format, randomly selected from the five different pattern types. Children were encouraged to explore how to solve the puzzles. This was done by first giving a short instruction and subsequently letting the child try the first task. If the child got stuck, a hint was given, for example, with the virtual version: 'How would you be able get this block in the puzzle?'. Once the child expressed that he or she fully understood the assignment, the other three puzzles were administered. The child was interviewed at the end of each version, to learn how the tasks had been experienced. Close observation of the child's behaviour was undertaken throughout the experiment.

3.2.4 Measures

The following metrics were used in the test: (1) Time: the task completion time for each puzzle was measured using a stopwatch; (2) Measuring and Fitting behaviour: the number of the child's measuring and fitting attempts was recorded; (3) Interview: the child was asked about how he or she had experienced the game, and about prior experience with using a computer; (4) Observations were recorded by the experimenter (nb. experiments were also recorded with a video camera).

3.3 Results

3.3.1 Task difficulty and overt problem solving

To answer the question regarding the comparative difficulty level of the tangible versus virtual task versions, completion times (in seconds) were compared. To normalise the distribution of scores, task completion times that were three standard deviations above or below the mean task completion time were excluded, and as a result, the data from one child were excluded from further analysis. Completion times for each of the two versions are shown in Table 3.3.

Table 3.3. Total task completion time in seconds (Mean and SD) per task version.

Version	Mean	SD
Tangible (n=25)	386.16	90.39
Virtual (n=25)	652.56	223.76

A paired-sample t-test on the total task completion time showed a significant result, t (24) = -6.72, $p < 0.01$, indicating that it took significantly more time to complete the four virtual puzzles than the four tangible ones. The results, presented in Table 3.3, revealed that children needed almost twice as long to solve the 4 virtual puzzles.

A possible effect of order of format administration was analyzed by using ANOVA with format order (version 1= tangible first; version 2= virtual first) as factor and the total solving time of 6 puzzles as the dependent variable. The first puzzles were left out of the analyses because these had been used to familiarize the children with the interface and the children had been helped to solve them. Significant effects for both solving time and order of task format administration were found: $F(1,23) = 29.21$, $p < .001$, $\eta 2 = .56$, and $F(1,23) = 4.88$; $p = .037$, $\eta 2 = .17$ respectively. This indicates that tasks were more difficult to complete in the virtual format than in the tangible format and that order of format administration had a significant influence on the total completion time of the virtual version, while it had not on the tangible version. The mean completion times for the tangible and virtual versions per condition are presented in Figure 3.2.

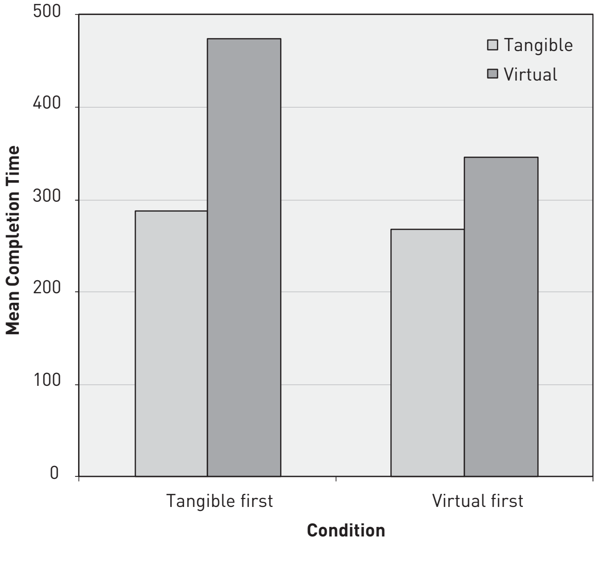


Figure 3.2. Mean task completion times per condition for the tangible and virtual version.

Analysing the relationship between scores on the virtual and tangible puzzle tasks by Spearman-Brown correlations revealed a considerable correlation ($r = .47$), indicating that both task versions measure a common component but measure unique aspects as well. Based on psychological task analysis, this is as expected.

Surprisingly, boys and girls did not significantly differ in the time it took to solve the tangible tasks ($p=.31$) but boys were quicker when the tasks were computerised ($F(1)= 6.8$, $p=.02$). Girls appeared to be more deliberate as they made 20% fewer fitting and measuring attempts with the tangible version and 9.4% fewer with the virtual version.

3.3.2 Practising

To study the influence that practice with the puzzles may have on children's performance, the average completion times for the second, third, and fourth puzzle were compared for both conditions. Since children were presented the puzzle types in a random order, the graphs in Figure 3.3 provide an indication of learning effects independent of task difficulty. Once again, the results from the first puzzles were excluded from the analysis because children were instructed on these tasks. We included the instruction tasks in Figure 3.3 though (dotted line), as it indicates how these tasks compared with the subsequent tasks. Repeated measures with the measurement intervals (6) and the two conditions as factors, and times taken to solve the tangible and virtual puzzles respectively showed a significant interaction between time and condition; $F(1,23) = 8.60$; $p < .001$; $\eta 2 = .27$. Regarding tangible puzzles, children in condition 1 showed different solving times between measurements 2 and 3 when compared with children in condition 2 ($p <.001$). Regarding virtual puzzles, differences between the conditions were found between measurements 7 and 8 but these did not reach statistical significance ($p = .06$). Again, it can be seen in Figure 3.3 that the completion of virtual puzzles took significantly more time than that of the tangible puzzles. Further inspection of the data revealed that the children in condition 1 (tangible first) needed on average more time to complete the virtual tasks.

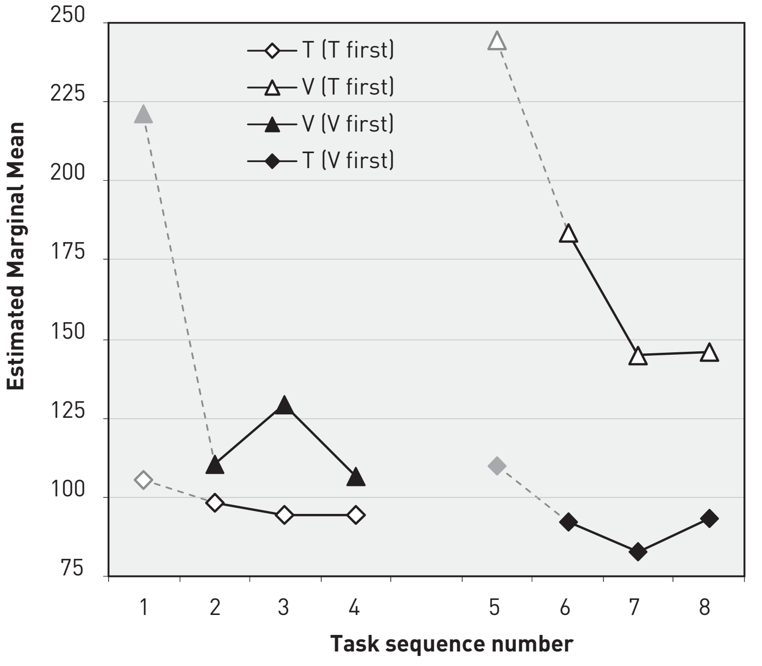


Figure 3.3. Mean task completion times for the tangible (T) and virtual (V) versions of the measurement instance sequences (condition 1 = T first, condition 2 = V first) including the instruction task (dotted line).

3.3.3 Overt problem solving behaviour

We analysed overt problem solving behaviour by comparing the number of 'measuring and fitting attempts' for both formats. A paired sample t-test on the number of 'measuring and fitting' attempts yielded a statistically significant result $t(24) = 3.321$, $p < .01$. Further inspection revealed that significantly more 'measuring and fitting' attempts were made in the tangible than the virtual version (see Table 3.4). Pearson correlation coefficients between number of measuring and fitting attempts and time for solving tangible puzzles ($r = .25$) and time for solving virtual puzzles ($r = .03$) confirm these findings.

Table 3.4. Means and standard deviations for number of measuring and fitting attempts for tangible and virtual version.

Version	Mean	SD
Tangible (n=25)	5.12	2.76
Virtual (n=25)	3.36	2.25

It can be concluded that significantly more 'measuring and fitting' attempts were made in the tangible version compared with the virtual version, but also that less time was needed when children solved the tasks with the tangible materials.

3.3.4 Children's opinions of both task versions

The children were asked what they thought of the puzzle tasks shortly after they had completed these. The questions addressed the extent to which they had found the tasks to be fun and how difficult they had been to solve. The first question concerned which version of the puzzle the children had found to be most fun to solve. Eight children answered that they had found the tangible version to be more fun than the virtual version; eight children could not choose between the two options, and nine children considered the virtual version to be more fun (see Table 3.5). The rationales that children gave for their preferences differed. Some preferred the tangible version because they found it easier while others preferred the virtual version, because they considered this latter one easier, in particular because the blocks were sorted already. Some children liked the difficulty of the virtual interface, presumably due to the fact that they needed to put more effort into the control of the PC. The tangible version may have been relatively easy for this group of children, which would explain why they appreciated the added challenge that the virtual version offered. Some children said that they preferred the virtual version simply because it was played on a PC.

Table 3.5. Interview answers in % for 25 children. 'T' means tangible, '-' means undecided, 'V' means virtual.

Question	T	-	V
More fun	32%	32%	36%
More difficult	28%	12%	60%
Would play again	48%	4%	48%

When asked which version of the puzzles was most difficult, seven children identified the tangible version, fifteen children found the virtual version and three children could not decide. This difference in perceived task complexity can be entirely attributed to the girls as all but two said that they had found the virtual version more difficult. This was the only question for which there was a clear gender bias.

The last question concerned which version of the puzzles they would choose if they were asked to solve one of the puzzles again. Twelve children chose the tangible version; twelve chose the virtual version and one child could not identify a preference. Thus, no clear preference emerged even though the virtual version had taken the children considerably more time.

3.4 Discussion

In this study, we considered whether complex visual-spatial reasoning tasks were easier to solve for children when administered in a tangible format than when presented on a personal computer as earlier research has not provided unequivocal answers (e.g., Olkun et al., 2005; Marshall, 2007). We have conducted an experiment with 5-7 year old children to investigate the influence that the task format can have on educational tasks in more detail.

The results of our study show that, overall, the tangible task set was solved almost twice as quickly as the virtual task set. This indicates that, of the two formats, the tangible version proved easier to use. In addition, we found that a greater amount of instruction was needed for the virtual version suggesting that this format was more difficult to use. A third finding is that the children showed different problem solving behaviour on the two task versions: with the tangible version we observed significantly more overt problem solving behaviour than with the virtual version. We will discuss these results below.

Our finding that the tangible task version is easier to use than the virtual version is supported by the significantly shorter task completion times for the tangible tasks. When we analyse both task versions, we see that the virtual version requires additional skills when placing the blocks onto the pattern. Since the virtual blocks were always offered in a horizontal orientation, the vertical blocks in the pattern required rotation and thus additional manipulation. This is reflected in the fact that there were longer task completion times for those puzzles that had a relatively large number of vertical blocks. The finding that the virtual task versions took significantly longer can be largely explained by the extra mouse manipulations needed to rotate blocks, which proved quite time consuming for many children. In fact, a simple model calculation revealed that in the virtual version the average completion time per task type largely depends on the number of rotations required to solve the puzzle rather than the total number of blocks. Therefore we may conclude that the task of using a mouse to rotate virtual blocks led to an extra challenge for most children. However, this type of challenge had not been intended for the educational goal of the Fitting and Measuring puzzle. The virtual tasks seem to require more "non-relevant" skills, for example, handling a mouse, for clicking and dragging actions for colour choosing and undertaking virtual block rotation. This suggests that a mouse may not always be a suitable attribute for children of this age group, an observation that has been suggested before (Hourcade, 2008). Mentally rotating objects, planning, fine motor skills, and extra working memory load are other factors that may have influenced children's performance on the virtual tasks. Screen-based tasks require rotation of elements in a two-dimensional space whereas they are representations of the three-dimensional wooden blocks. In a sense, this is unnatural and requires extra 'spatial translation' and memory load.

The first puzzle that children solved in each task version was for practice. The intention was that children should subsequently understand the interface and the goal of the task. Although most children seemed to understand the nature of the task and the use of the interface after the first puzzle, and all of them had had prior computer experience, in the virtual format, a number still found it difficult to rotate the blocks and drag them onto the pattern. For this reason, more help was needed primarily to help the children learn how to use the mouse to 'drag' the blocks onto the pattern. In contrast, hardly any additional instruction or guidance was needed with the tangible tasks as children readily understood this format. As a result, we found that the practice virtual task took significantly longer than the practice tangible ones. The latter tasks showed only a very small decrease in task completion time when the second set of tangible tasks was undertaken. This indicates that children quickly learned how to do the tangible task and could not become much faster. With the virtual version, children became much faster when moving from the practice to the subsequent tasks.

We found that children showed more overt problem solving behaviour with the tangible version, as significantly more 'measuring and fitting' attempts were made compared to the virtual version. For example, children would try to fit a block in the puzzle and then realize that it was too short, which prompted them to put this block back and select the correct one. Even though the tangible tasks were completed with more overt visual-spatial reasoning behaviour, they were still solved more quickly than the virtual ones. This could imply that the tangible puzzle version invites the child to show more of the necessary exploratory behaviour or a more systematic strategic approach. However, it could also imply that a task with tangible materials, such as the puzzle-task used in our study, requires different cognitive and/or motor skills than a similar looking task administered through a different interface. In the tangible version, children are presumably able to see immediately what happens as a consequence of try-and-measure behaviour and to react to this by direct movements which are likely to facilitate their further reasoning, comparing and matching behaviour. The tangible tasks can be assumed to appeal to the use of skills that are readily available. Furthermore, the tasks used in the study are similar to those that the children use in the classroom.

Solving the puzzles we used in this study required visual pattern analysis, visual comparison, reasoning and rotation. Although it was expected that girls would need more time to solve both tangible and computer puzzles (e.g., Levine, Huttenlocher, Taylor, & Langrock, 1999; Voyer, Voyer, & Bryden, 1995), they only needed more time to solve the virtual puzzles. This result can be best explained by the fact that most of the boys were more experienced in the use of a PC than the girls. Interestingly, girls were more deliberative when solving both item types. PC skills also seem to strongly influence children's preference for one of both interfaces. Some children preferred the virtual task condition because it was more challenging, others found it too challenging and preferred the tangible interface.

Although the study included a relatively small group of participants (N=25), we believe that the results are clear and should be taken into account when developing educational tasks, especially for the age group that we studied. Even though children are increasingly experienced in working with PC's, it seems important to realise that the requisite skills, e.g. to control the mouse, cannot be assumed at this age. We conclude that a virtual task format is less suited for an educational task such as the one we used in our study than a tangible task format. The virtual task format led to specific problems which detract from the task itself. In addition to the task goal, the children have to learn to cope with the challenges added by the format in which the task is presented. This is less than optimal, unless handling the interface is a separate and explicit educational goal. Therefore, the danger that a virtual interface may prove less effective for some learning and assessment situations should be taken seriously.

The combination of shorter solving times and at the same time making more use of "measuring and fitting" strategies forms a strong basis for our conclusion that, for the group of 5-7 year old children, at least in the complex visual-spatial reasoning domain, tangible educational tasks are preferable over screen-based ones.

In addition to the benefits of tangible interfaces, described in this paper, there are good reasons to use computers to support education. In classroom environments teachers wish to know how individual learning progresses over time, and training and assessment through a computer can be a welcome addition to current teaching practice. In educational assessment, the use of electronic tangibles may prove valuable by offering specific information such as how much time it took to move one object to another location or which object was taken first. On a higher level, errors can be captured and examined, and in the future these may be categorized and used for dynamically training the child. Furthermore, applied strategies may be identified from the way objects are manipulated during a task. The information stored on the computer can be used to adapt tasks to the current abilities of the child or to his or her development within their zone of proximal development (e.g., Vygotsky, 1978). This would help to ensure that tasks presented are neither too difficult nor too easy, thus providing an optimal condition for learning and preventing the child from becoming frustrated or losing interest. Of course, this requires further and finer cognitive task and process analyses for all tasks individually (e.g., Resing, 2006a, b). Adaptive training cannot become optimal without a thorough analysis of psychological processes involved in the solution of the (complex) cognitive tasks and the object manipulations of the child who solves the task. This includes task-analysis and further research concerning dynamic ways of scaffolding learning so as to ensure that, with these tasks, the intended skills or abilities can be followed, enhanced, trained, and measured in the most effective way.

Future research will be focused on extending the use of tangible electronics to tasks similar to those that we used in this study, and also to more complex tasks that can cover a wider range of skill development. This requires investigation of the use of tangible electronics in school materials and integrating physical objects with electronics. Initial prototypes have been developed at a number of research institutes (e.g., Fontijn & Mendels, 2005; Khandelwal, 2006; Verhaegh et al., 2007) and a number of evaluations have been carried out (Dijksma, Resing, Roig, Van Herk, Meijles, & Verhaegh; 2008; Resing & Elliott, 2008; Roig, Resing, Van Herk, Meijles, & Verhaegh, 2008). In the near future, extensive research will be needed in order to ensure that maximum benefit can be gained from the combined positive aspects of tangible educational materials and computing technology.

4

Study II

On the correlation between children's performances on electronic board tasks and nonverbal intelligence test measures

Verhaegh, J., Fontijn, W.F.J., & Resing, W.C.M. (submitted). On the correlation between children's performances on electronic board tasks and nonverbal intelligence test measures.

Abstract

In this study it was investigated whether a tangible electronic console (TagTiles) can be used to address a range of cognitive skills. The tasks implemented on the console consisted of abstract visual patterns, which were intended to target perception, spatial knowledge representation, eye-hand coordination, reasoning and problem solving. The results of a pilot study (N=10, children aged 8-10) and an experiment (N=32, children aged 8-10) are presented. Correlations between scores on TagTiles tasks on the one hand and a selection of WISC-III^{NL} performance subtests, Raven's progressive matrices and RAKIT's Memory Span on the other hand, were calculated. The results indicate that the TagTiles tasks cover similar skills as the applied WISC-III^{NL} subtests, demonstrated by the moderate to large correlations between performance scores on sets of TagTiles tasks and sets of WISC-III^{NL} tasks. The combined TagTiles task scores were also significantly correlated with the aggregated WISC-III^{NL} subtest scores. Significant correlations were found between the TagTiles tasks and the Raven test scores, though for the RAKIT Memory Span no significant correlation with TagTiles tasks was found. After further refinement and validation the tasks can be applied to provide an indication of children's skill levels, offering the benefits of a self motivating testing method to children, and avoiding inconsistencies in administration. As such, the tasks may become an effective tool for the training and assessment of nonverbal skills for children.

4.1 Introduction

Conventional psychological tests are important tools for psychologists and often far reaching conclusions are based on their outcome. Consequently, before a test is trusted and becomes widely applied, it is usually validated with a large number of individuals (e.g. Wechsler, 1949). Thus, systematic errors can be detected and corrected for and noise can be averaged out. Special committees (such as COTAN for Dutch tests; Evers, Lucassen, Meijer & Sijtsma, 2010) evaluate tests on several criteria such as reliability, validity and quality of the test materials before they become available for use as a psychological instrument. However, even if we assume that it is established that a test assesses the skills that it is intended for, some problems remain. In the present article we describe a new tool which may help to overcome a number of these problems, by offering a test-solution that can be seen as a 'light-weight' version of existing tests, in the sense that it delivers a less formal screening of abilities, though the test is easier to take and to administer. We present an empirical study in which we tested the validity of an application of the tangible electronic console 'TagTiles' (see Figure 4.1) for testing and training purposes. TagTiles is a tabletop electronic console with tangible playing pieces, developed by Serious Toys B.V. (www.serioustoys.com). The console includes a tabletop sensing board with an array of LED lights underneath and audio output via separate speakers. Below we describe a number of issues with current test practice. Some of these may be mitigated to a certain degree by applying TagTiles.

Figure 4.1. TagTiles console with an example application to train verbalizations of spatial relations. The playing pieces are localized and identified by the system. Visual effects are created with LED lights underneath the surface (top right square emits green light in this example).

Circumstances during the administration of a test can introduce systematic and random errors due to individual differences between the individuals being tested and/or the person administrating the test. These individual differences, such as a participant's mood at the moment of testing, are inherently unpredictable and therefore it is impossible to fully correct for them or to filter out the possible effect they have on the outcome. Also, intelligence tests are rather rigid in their 'judgment', as they often provide an IQ score based on a single assessment (e.g., Sternberg & Grigorenko, 2002; Elliott, Grigorenko & Resing, 2010). The same test may lead to different results due to different circumstances.

We discuss a number of issues that can affect standardized assessments, in particular with young children. These include anxiety and stress that a test may cause in a child; children may lack the motivation to participate or do well on a test; irregularities in the test administration and scoring can occur; not all relevant skills can be assessed properly with existing tools (e.g. motor skills); the test format (e.g. pen/paper, PC) may offer an added challenge that varies per child. Some of these issues are more important than others, though all of them may introduce 'noise' in the result or even cause the result to be wrong.

The first two issues relate to the individual response of a child to a test situation; so-called non-intellective factors (Tzuriel, 2001). Many children feel stressed by the idea of being tested (McDonald, 2001), leading to test results that underestimate their true capabilities (Hembree, 1988; Appl, 2000; Thurman & McGrath, 2008; Meijer, 1996). This is especially problematic if the test is used as a diagnostic tool to determine the developmental level and possibly reveal a developmental lag. Next to or other than experiencing stress, children are often not properly motivated to participate in a test (Duckworth, Quinn, Lynam, Loeber & Stouthamer-Loeber, 2010), for example because they dislike being taken out of the classroom to take a test that other children do not have to take, or, perhaps more importantly, because they find the test itself boring.

The third issue relates to the role of the test leader. In most tests, a test leader is needed to provide instructions to the child. Although test manuals provide elaborate instructions on how to administer the test to the child, differences in the way these instructions are executed will occur, for example due to misunderstanding of the instructions or lack of experience, both leading to errors in the administration procedure (e.g. Kuentzel, Hetterscheidt & Barnett, 2011). Also differences between test leaders in their style, motivation, character or mood can influence the test outcome. In addition, recording the test responses is usually done manually by the test leader. Again this may lead to variability, as children's responses in many cases need to be judged by the test leader immediately during the test. A similar response may be judged differently by different test leaders or even by the same test leader at different times. As a result, mistakes and/or missing values will occur in scoring records.

A fourth issue is that, due to their interactive nature, paper and pencil-based tests cannot assess the full spectrum of relevant skills. Especially executive functions, such as attention and (working) memory are difficult to assess in a structured manner without a PC-based task. However, a number of studies has shown (e.g. Bayliss, Jarrold, Gunn & Baddeley, 2003; Cowan, Saults, & Elliott, 2002; Swanson, 2006) that these skills are crucial for children's ability to become skilled at others, such as literacy skills and mathematics. Working memory was found to be a more powerful predictor of subsequent academic success than IQ during the early years (Alloway & Alloway, 2010). Early screening and intervention should therefore be based on the assessment of working memory and executive functions as well. Also motor skills are only addressed minimally in conventional IQ tests for children older than 6 years, even though they were found to contribute to children's academic achievements (Grissmer, Grimm, Aiyer, Murrah & Steele, 2010). Currently, fine motor skills are assessed mostly by judging precision and accuracy in writing (e.g. the Symbol test in WISC-III^{NL}). Adding interactive properties to the testing tools, will allow for testing motor skills more extensively (e.g. planning movements) and more precisely, which is impossible with non-interactive tests. Children can be triggered to make specific motions and also these motions may be recorded and analyzed automatically with a tangible electronic interface (e.g. as applied in a therapeutic context by Li, Fontijn & Markopoulos, 2008).

Part of the problems described above, can be solved by using PC-based tests. Indeed, in addition to tests offered in pen-and-paper format, PC-based tests have been introduced over the last years, such as the Cambridge Neuropsychological Test Automated Battery (CANTAB®, e.g. Lehto, Juujarvi , Kooistra & Pulkkinen, 2003), and also digital training programs such as 'Jungle Memory' (www.junglememory.com) are available. Using a computer to provide instructions to the child guarantees identical instructions in each test, provided in an identical manner. However, for computer (semi) illiterate users, e.g. young children, the computer interface is a problem (Hourcade, 2008; Verhaegh, Resing, Jacobs, & Fontijn, 2009). Furthermore, as explained above, only a part of the relevant skills can be addressed properly with a PC and controlling the computer adds an extra challenge. Physical electronic interfaces can be applied to address motor skills directly, allowing direct manipulation of objects, without interference by an interface that may add an unwanted challenge to a test, such as manipulating a computer mouse (Verhaegh et al., 2009). Absence of readily available technology to enable test formats with a physical interface may be the reason why the inclusion of assessment of motor skills is limited in current tests.

The aim of our experiment was to validate whether a newly developed set of tasks for the physical electronic interface TagTiles is suitable to address a range of nonverbal skills, as a precursor to being able to measure these skills. To this end, the TagTiles task scores of children were correlated with their performances on several non-verbal conventional psychological tests. This experiment is part of a larger study with the aim to use fun as an effective instrument in enhancing children's intrinsic motivation to give their best performance at a task, without being hampered by performance anxiety or boredom.

Introduction of the TagTiles tasks

Eight different tasks were developed for the experiment. The tasks are based on abstract colored patterns of squares that fit an 8x8 grid (see Table 4.1). Different operations need to be performed on the patterns, depending on the task. Each task consists of a number of items of increasing difficulty, and each item contains one pattern. The tasks were created to address a diversity of visual-spatial skills and they are based on the tasks that have been developed for an earlier explorative study (Verhaegh, Fontijn & Hoonhout, 2007).

In Table 4.1 a description is provided of the tasks that were developed, illustrated with patterns of level 2 of each task. The eight tasks can be subdivided into three skill categories which we labeled basic skills, such as attention and perception; memory skills, including active and passive memory tasks, and complex skills related to executive functioning such as spatial reasoning and problem solving. In the following paragraphs we describe how we expected our tasks to fit in the three skill categories.

Tasks 1 and 5 (see Table 4.1) were expected to address mostly basic skills. Task 1, tagging the pattern, is straightforward and does not require much cognitive processing. In Task 5, selective attention is combined with these basic skills: the squares with blinking lights need to be ignored, requiring inhibition of (motor) responses. Task 1 was meant to introduce the child to the task interface and Task 5 was meant to give a break from the more complex tasks that are preceding and succeeding. This break was inserted to maintain or restore the child's motivation. Basic skills are also assumed to be required for the other six tasks, but in Tasks 1 and 5 they are addressed in a more isolated manner.

Memory skills are aimed to be addressed by Tasks 2, 3 and 4 (see Table 4.1). In Task 2 a sequence of squares needs to be reproduced in the exact same order as the assignment pattern, but since the location of the pattern on the board is indicated in the task, there is no need to memorize the spatial location of the pattern. However, spatial sequential memory is required as the colors of the squares and their order has to be reproduced. In Task 3 working memory load is added as the sequence needs to be reproduced in reversed order. In Task 4 the full pattern and its spatial location have to be reproduced, requiring passive spatial memory. In Task 6 the pattern is again presented as a sequence, but because the task is to extrapolate the sequence instead of reproducing it, reasoning skills are combined with spatial sequential memory. Memorization of the sequence helps to identify the rule that is underlying the pattern, and thus helps to finish the pattern.

More complex spatial skills were targeted with Tasks 7 and 8, and also with Task 6 in the higher levels. In Task 7, a pattern needs to be mirrored to another location on the board, requiring spatial reasoning and knowledge representation. Task 8 is similar, though the pattern needs to be mirrored using a diagonal axis (see Figure 4.2, right picture), which requires more complex spatial reasoning than using the horizontal axis in Task 7.

Table 4.1. Description of the eight TagTiles tasks, including targeted skills as anticipated during development. Displayed patterns are taken from the experiment (from each task level 2 is shown). Targeted skills are written in order of importance for the task. Printed in bold are the skills we expected to be most relevant per task. Numbers (not displayed in the task) indicate sequence of appearance, striped shading indicates blinking, grey shading and crosses indicate the correct solution.

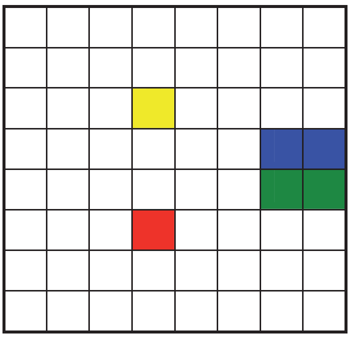
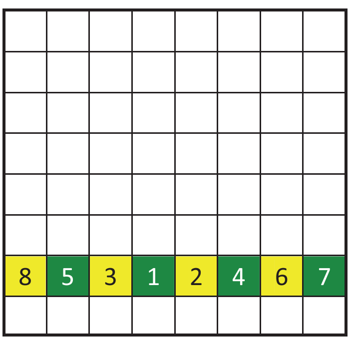
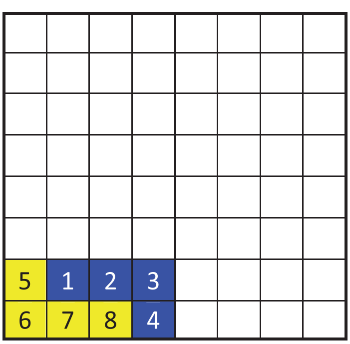
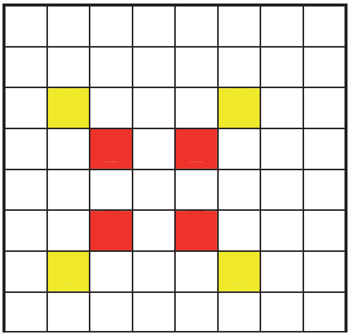
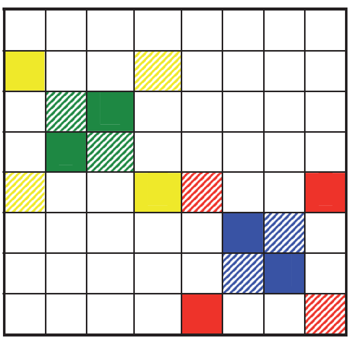
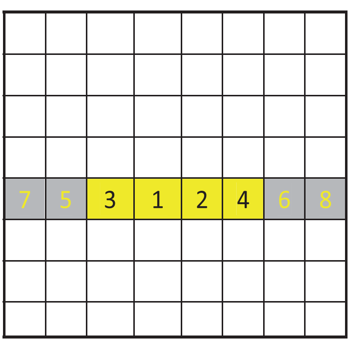
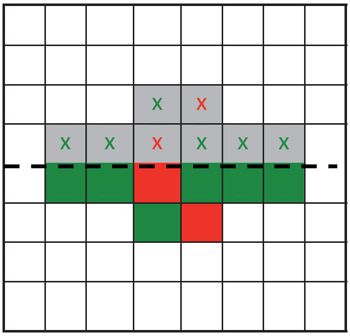
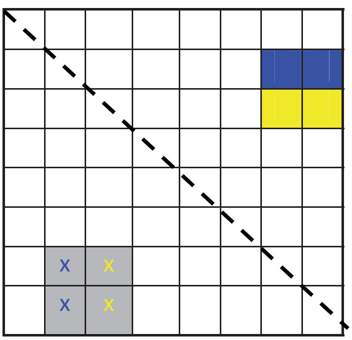
Example patterns	Task	Targeted skills
	**1. Tag the pattern**	
	A pattern is displayed and the task is to tag the pattern on the board with pawns with the corresponding colors. The pattern is displayed on the LEDboard continuously throughout the task.	**Fine motor skills, eye-hand coordination and perception,** attention skills, spatial memory, higher levels also require reasoning.
	**2. Reproduce the sequence**	
	The pattern is built-up from a sequence of colored squares. Once complete, the pattern turns white and a sound signal indicates that the player may start. The white pattern is continuously displayed until the colored pattern is reproduced in the correct sequence.	**Knowledge representation, spatial sequential memory,** attention, fine motor skills, eye-hand coordination and perception, (passive) spatial memory.
	**3. Reproduce the sequence in reversed order**	
	This task is largely identical to task two, except for the difference that the sequence has to be reproduced in reversed order.	**Knowledge representation, spatial sequential memory,** attention, fine motor skills, eye-hand coordination and perception, (passive) spatial memory.

Table 4.1. Continued.

Example patterns	Task	Targeted skills
	**4. Reproduce the pattern from memory** A pattern is displayed on the LEDboard. After a few seconds, the pattern is taken away, followed by a short delay ending with a sound signal, which indicates that the pattern should be.	**Spatial memory (passive), knowledge representation,** attention, fine motor skills, eye-hand coordination and perception.
	**5. Tag the pattern, ignore the blinking squares** A pattern appears on the LEDboard, containing squares that blink (distractors) and squares that do not. The task is to tag the non-blinking squares.	**Attention, fine motor skills, eye-hand coordination and perception,** (passive) spatial memory.
	**6. Extrapolate the pattern in a logical manner** A pattern is built-up in a sequence. The pattern on the LEDboard is displayed continuously and the task is to add a sequence with the same length as the example pattern, which would follow logically from the example pattern.	**Reasoning, knowledge representation, spatial sequential memory,** attention, fine motor skills, eye-hand coordination and perception.

Example patterns	Task	Targeted skills
	**7. Mirror task horizontal** The task is to mirror the pattern that is displayed. The board is split in two parts by a dotted horizontal line that serves as the mirror line.	**Knowledge representation, spatial reasoning,** fine motor skills, eye-hand coordination and perception, attention, (passive) spatial memory.
	**8. Mirror task diagonal** Again the task is to mirror the pattern, but this time using a diagonal mirror line which is displayed on the board.	**Spatial reasoning, knowledge representation,** fine motor skills, eye-hand coordination and perception, attention, (passive) spatial memory.

We used a set of rules to initially estimate the relative complexity of each pattern. The rules we applied are the following: (1) (a)symmetry of a pattern, (2) the number of colors used, (3) the size and (4) the location of the pattern on the board (adjacency to a border makes a pattern easier to memorize; the border forms a landmark). The composition of levels per task was made in such a way that we expected only few children in the target group (8-10) to be able to complete the highest levels, but all should be able to complete at least the first four levels. The reasoning behind this was that each task should offer enough challenge for all children within the target group, but at the same time be easy enough to achieve some progress.

In terms of cognitive processes that were expected to be targeted with the TagTiles tasks, our study can be related to research that has been done regarding children's working memory by Mammarella and colleagues (Mammarella, Pazzaglia & Cornoldi, 2008). Working memory (WM) is claimed to be a fundamental skill in learning (Alloway & Alloway, 2010) and therefore tools to assess WM are needed. Mammarella et al. have applied a test battery for a broad assessment of different visual-spatial working memory mechanisms. This battery included Visual Pattern tasks to address active (complex span) and passive (simple span) spatial working memory, quite similar to the pattern tasks included in our study. Their study provided evidence for the existence of different visual-spatial working memory components, and thus the relevance of testing tools addressing these components. In their study they tested different theoretical models of working memory. The best fitting model for the WM task data was found with the 'continuity'-model which includes three continuous dimensions of WM: visual WM, simultaneous-spatial WM and sequential-spatial WM (Mammarella et al., 2008). Because of the similarities amongst the pattern tasks applied in their study and the TagTiles tasks, their results provide some indication that the TagTiles tasks can also be used to assess the three mentioned dimensions of WM.

Since the TagTiles tasks were expected to address other abilities besides WM, and the number of tests included in our experiment had to be limited, we chose to compare TagTiles with tests that address a broad set of nonverbal IQ related skills. Scores on the TagTiles tasks on the one hand were correlated with scores on a selection of WISC-III^{NL} subtests (Wechsler, 2005), RAKIT (Bleichrodt, Drenth, Zaal & Resing, 1984) and Raven's progressive matrices (Raven, Raven & Court, 2004) on the other hand. It was hypothesized that there would be significant correlations amongst the performances of children on the TagTiles tasks and the applied non-verbal conventional psychological tests. Specific expectations regarding correlations are presented in Table 4.2.

Table 4.2. The expected correlations amongst the TagTiles Tasks (T1-8) and the conventional psychological tests, and their mutual skills.

	Measure	**TagTiles Task**								**Assumed mutual skills**
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	
		Tag pattern	*Reproduce sequence*	*Reproduce sequence reversed*	*Reproduce memorized pattern*	*Tag ignore blinking*	*Extrapolate pattern*	*Mirror pattern horizontal*	*Mirror pattern diagonal*	
WISC-III^NL	*Picture Completion*	**	*	*	*	*	*	*	*	Perception, eye-hand coordination, motor skills
WISC-III^NL	*Coding*	**	*	*	*	***	*	*	*	(Selective) attention, perception, eye-hand coordination, motor skills, visual memory
WISC-III^NL	*Block Design*	**	**	**	**	**	**	***	***	Perception, spatial knowledge representation, eye-hand coordination, reasoning and problem solving with geometrical patterns, motor skills
WISC-III^NL	*Object Assembly*	*	*	*	*	*	**	***	***	Perception, spatial knowledge representation, eye-hand coordination, reasoning and problem solving with everyday objects, motor skills
WISC-III^NL	*Symbol Search*	**	*	*	*	***	*	*	*	Visual memory, (selective) attention, perception, eye-hand coordination, motor skills
WISC-III^NL	*Digit Span*	*	**	**	**	*	*	*	*	Attention, sequential memory, though TagTiles targets *visual spatial* memory and Digit Span measures *verbal* sequential memory
WISC-III^NL	*Vocabulary*	-	-	-	-	-	-	-	-	None
	Standard Progressive Matrices	*	*	*	*	*	***	***	***	Spatial reasoning and pattern completion
RAKIT	*Memory Span*	-	**	**	**	-	-	-	-	Sequential memory, though TagTiles targets *spatial* sequential memory and Memory Span measures *verbal* sequential memory

Note. For '***' we expected to find significant Pearson correlation coefficients of $r > .5$; for '**' >.3 and for '*' >0.1 (values based on Field, 2009, p. 170).

4.2 Method

4.2.1 Pilot test

A pilot test was conducted with ten participants, four boys and six girls with a mean age of 9 (111 months, SD 8.7 months). Six of the tasks described in Table 4.1 were administered, omitting Task 3 and Tasks 7 and 8 were combined in one task in the pilot test. Children were tested individually at a test facility of Philips Research, called the 'Experience Labs', which is a laboratory that offers a comfortable homelike environment.

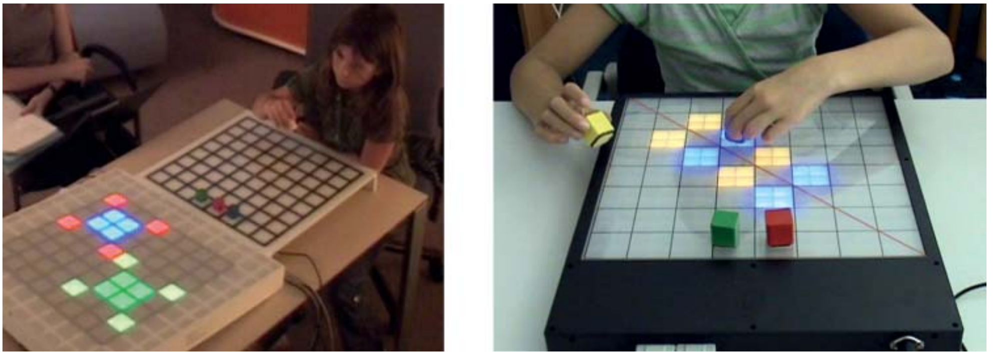


Figure 4.2. TagTiles. Left: A participant completing a pattern tagging level (Task 1) in the pilot study. Right: A participant completing one of the mirror patterns (Task 8) in the experiment.

To administer the tasks two TagTiles boards were used: one for system input from the children provided by tagging the board with the blocks, and attached one board to display the assignments (see Figure 4.2, left picture). The results from the pilot test showed that the levels within tasks yielded gradually increasing level completion times for most tasks. The average level completion time was used as a measure of difficulty, as it includes thinking time, task execution time, and time lost due to committed errors. Also we observed that transferring the example pattern from one board to another one, often yielded an extra challenge. We learned which instructions were most effective in explaining the tasks to the children. The results of the pilot test were used to optimize the set up of the subsequent experiment. For the experiment it was decided to use a new prototype of the TagTiles board (Figure 4.2, right picture) that only included one board with integrated system input and output options.

4.2.2 Experiment

Participants

Thirty-two children participated in the experiment. They were all in the fourth grade ('groep 5' in the Dutch school system) of a regular primary school in the south of the Netherlands. Twenty-three boys and nine girls participated, aged 8-10 years (mean age of 113 months, SD= 11.5 months). Informed parental consent was obtained for all children prior to the start of the experiment. All participants completed a short version of Ishihari's test for color deficiency successfully. Each participant received a 'TagTiles-certificate' and some cards (e.g. with Pokémon) as a reward for their participation.

Design

Each child participated in three test sessions which were spread over 6 weeks time. The first session included an individual test with the WISC-III^{NL} subtests. In the second session Raven's progressive matrices were administered in groups (about 10 children per group). The third session consisted of Ishihari's test for color deficiency, the TagTiles tasks and the RAKIT visual-spatial memory task administered individually. The sessions took place in a quiet room in the school. Each session lasted one hour or less.

Measures

WISC-III^{NL} subtests

A selection of performance subtests of the Dutch version of the Wechsler Intelligence Scale for Children, third edition NL (Wechsler, 2005) was administered to all participants. The following subtests were included: Coding, Symbol search, Picture completion, Block design, Object assembly, Digit span and Vocabulary. Coding and Symbol Search together are assumed to assess the WISC-III^{NL} factor Processing Speed including visual motor coordination, reflection, visual memory, working in an orderly manner, motivation and fast processing of visual information. Picture Completion, Block Design and Object Assembly are claimed to make up the WISC factor Perceptual Organisation assessing attention for details, analyzing patterns and combining parts to a whole. Digit Span is not seen as part of the WISC three-factor model; it has shown weak correlations with total IQ as well as other subtests (Wechsler, 2005). The Vocabulary subtest was administered to include an assessment of skills that is not related to spatial perception or short term/working memory, as part of the verbal WISC factor.

Raven

Raven's Standard Progressive Matrices (Raven, Raven & Court, 2004) were administered in pencil-and -paper format to assess IQ from a pattern completion and reasoning perspective. Raven's test was chosen as it is also a non-verbal test (as are the TagTiles tasks) and it appeals on deductive reasoning skills with patterns, which might be needed for TagTiles as well. Instructions were provided to the children as prescribed by the Raven manual.

RAKIT subtest

We used part of the Memory span subtest of the RAKIT (Revisie Amsterdamse Kinder Intelligentie Test, Bleichrodt, Drenth, Zaal, & Resing, 1987) to assess sequential visual memory for concrete figures. A sequence of figures (tree, chair, etcetera) was shown on paper and then hidden, and the child needed to reproduce the sequence with cards.

TagTiles tasks

The TagTiles tasks were administered on the TagTiles console with a single board (see Figure 4.2, right picture). The console was positioned on a table in front of the child. All eight tasks as described in Table 4.1 were administered. At the start of the test session it was explained to the child that there were eight different 'game' tasks (to prime the expectation of a game) and that before the start of each, the experimenter would explain the task. The child was asked to indicate whether the task explanation was understood; if not, the experimenter would explain again. There was one demonstration level and two practice levels. When the child demonstrated to the experimenter that he or she had understood the task after completing one practice level, a second practice level was skipped. No help or hints were provided after the child had started the task. To complete a level, children interacted with the board by tagging squares using four colored wooden blocks (blue, red, green and yellow) that were placed at the top end of the board (from the child's perspective, see Figure 4.2). The blocks needed to be placed back after each level to ensure precise measurement of task completion time and to prevent children from taking blocks in their hands before the start of the level. Children directly manipulated the blocks on the board, while the system registered each block manipulation. Each level was accompanied by sound signals to indicate the start of a level, and also when touching the board with a block, auditory feedback was provided to indicate whether the 'move' was right or wrong. After each successful completed level an 'applause'-sound is heard. We followed a typical span self-terminating procedure, where the task was stopped when the child could not complete two subsequent levels within 90 seconds each. However, when the child was close to the solution at 90s, the time was extended to 110s maximum, to avoid discouragement.

Scoring and analysis

For the WISC-III^{NL} subtests we used the raw score data for analysis, as the participants were in a narrow age range (8-10 years). For TagTiles a method was determined to calculate scores. These scores could be computed based on three measures: the highest level completed by the child, the level completion time, or the number of levels that was completed within a level-time threshold. The first measure was considered not appropriate because some levels were not completed, but subsequent levels were. Using the number of levels that was completed within different time thresholds (e.g. 15, 30 and 60 seconds) and relating this to a score would be an insensitive method, as part of the variance in scores would be ignored. Therefore we chose to use the second measure, using the level completion times as 'scores'; as level time also includes thinking time and the time (to correct) errors takes. A level score was calculated by subtracting the actual level completion time from the level cut-off time (110 seconds for all levels). We used this cut-off time because typically when children had not finished the level by that time they would not finish at all. This means that in case it took 110 seconds or longer to complete the level, the level was considered to be 'failed' and this led to a level score of 0. Children who failed a level for other reasons, e.g. because they gave up, also received a level score of 0.

Pearson correlations were calculated among the individual TagTiles task scores and the subtest scores of the conventional tests. We then grouped scores of WISC-III^{NL} subtests on the one hand and TagTiles tasks on the other hand to examine correlations among them. The TagTiles scores were grouped as described in Section 2: tasks 1, 5 (basic skills), tasks 2, 3, 4 (memory skills), and tasks 6, 7, 8 (complex skills). The scores of administered WISC-III^{NL} subtests were grouped according to the three WISC factors: Verbal Comprehension, Perceptual Organisation and Processing Speed (Wechsler, 2005).

To examine the relationship between the TagTiles scores and the test scores we used the Partial Least Squares (PLS) regression method (Smilde, Bro & Geladi, 2004) to predict the scores on the WISC-III^{NL} with the TagTiles scores, and to find the number of underlying factors. PLS is considered appropriate to predict responses when there are many factors that are highly collinear (Tobias, 1995). Applying a Principal Component Analysis (PCA) was considered less optimal as PCA summarizes the TagTiles scores independently of the scores on conventional tests, whereas PLS summarizes the TagTiles scores in factors that correlate best with these tests. Also the total WISC-III^{NL} nonverbal subtest scores were correlated with TagTiles total scores. The different TagTiles levels were analysed for their difficulty using completion times with ANOVA to investigate how different pattern properties contributed to the complexity of a pattern.

4.3 Results

Descriptive statistics for the mean scores obtained on the conventional psychological (sub)tests and the TagTiles tasks are presented in Table 4.3. Raw data are presented, since there are no norms for TagTiles yet. The data for the conventional (sub)tests are within the average range for the respective IQ (sub)test scores when compared with the ranges published in the test manuals. Some TagTiles tasks showed relatively large ranges in level completion times (larger than 40 s), indicating that these include a wider range in difficulty level than others with smaller SDs. Strikingly, Task 3 showed a significantly higher mean completion time and a smaller SD than Task 4.

Table 4.3. Descriptive statistics of TagTiles scores and raw scores on conventional tests obtained by 32 participants.

			Range	
Test	M	SD	Min	Max
TagTiles (seconds)				
1. Tag the pattern	14.00	2.17	10.72	20.55
2. Reproduce the sequence	49.59	13.00	23.77	83.45
3. Reproduce the sequence backwards	34.12	9.88	13.57	49.49
4. Reproduce the pattern from memory	28.24	15.42	9.37	69.96
5. Tag the pattern but no blinking squares	17.22	3.44	12.18	26.57
6. Finish the pattern in a logical way	20.86	9.32	5.73	43.44
7. Mirror task (horizontal)	25.17	5.91	16.86	39.36
8. Mirror task (diagonal)	40.46	12.82	18.15	63.60
Tagtiles Total score	28.71	4.75	20.82	36.92
WISC-III^{NL} (raw scores)				
Picture completion	19.34	2.27	14.00	24.00
Coding	36.75	8.52	19.00	53.00
Block Design	38.47	17.21	21.00	58.00
Object assembly	27.09	5.68	15.00	36.00
Symbol Search	21.78	3.16	14.00	30.00
Digit Span				
Forward	7.84	1.55	5.00	10.00
Backward	4.16	1.32	2.00	7.00
Vocabulary	28.34	5.45	16.00	39.00
RAKIT (raw scores)				
Memory span	10.19	2.20	5.00	14.00
Raven's progressive matrices (raw scores)				
Total score	35.84	6.77	20.00	55.00

The correlation coefficients among the TagTiles task scores and the psychometric test scores are shown in Table 4.4. The expected medium to large correlations (as presented in Table 4.2) between TagTiles tasks and conventional psychological tests are largely present in the data. The largest number of significant correlations was found with the subtest Block Design: four at a p 0.01 level (2-tailed). Correlations with Object Assembly performances showed a similar pattern, though with smaller correlations than with Block Design, and no significant correlations with Tasks 2, 3 and 7 were found. Tasks 1 and 2 showed large correlations with Symbol Search, but for Task 5 and other tasks only a medium correlation was found with this subtest. No significant correlations were found with Digit Span Forward, except for Task 8 where we found a significant, though negative, correlation. Similar to Digit Span, the TagTiles tasks showed no significant correlations with RAKIT Memory Span. The average scores of the TagTiles tasks showed significant correlations with three WISC-III^{NL} subtests: Block Design, Object Assembly and Symbol Search; the largest correlation coefficient was found with Block Design, r=.62. No significant correlations were found with Vocabulary, as predicted.

For Task 6 a significant correlation (two-tailed, $p < .01$) with Raven's scores was found. Task 1 of TagTiles showed a small correlation with Raven, though for other TagTiles tasks we found moderate correlations. Also the average total TagTiles score shows a significant correlation (.55 at $p<0.1$) with Raven.

The correlation coefficients in Table 4.4 suggest that there is a strong overlap between skills needed for the TagTiles tasks and the applied conventional test measures. We continued our analyses by calculating correlation coefficients among clustered scores.

Table 4.4. Correlations among the TagTiles Tasks and conventional psychological measures (N= 32)

	Measure	TagTiles Task								
		1 *Tag pattern*	**2** *Reproduce sequence*	**3** *Reproduce sequence reversed*	**4** *Reproduce memorized pattern*	**5** *Tag ignore blinking*	**6** *Extrapolate pattern*	**7** *Mirror pattern horizontal*	**8** *Mirror pattern diagonal*	*TT Average*
WISC-III^NL^	*Picture Completion*	-.18	.23	.08	.17	.10	.00	.24	.05	.16
	Coding	.27	.16	.23	.16	.12	-.11	.13	.24	.21
	Block Design	.08	.37*	.38*	.58**	.45**	.31	.46**	.47**	.62**
	Object Assembly	.08	.13	.14	.36*	.37*	.26	.33	.45**	.42*
	Symbol Search	.47**	.51**	.04*	.41*	.38*	.03	.35*	.02	.42*
	Performance Total	.22	.38*	.39*	.51**	.43*	.18	.44*	.45**	.57**
	Digit Span Forward	.01	-.33	-.09	-.13	-.06	.12	-.20	-.49**	-.28
	Digit Span Backward	.05	.06	.06	.00	-.02	-.08	-.06	.22	.07
	Digit Span Total	.04	-.19	-.03	-.09	-.06	.03	-.20	-.23	-.16
	Vocabulary	-.01	.29	.17	.25	.29	.13	.22	-.02	.23
Raven	*Standard Progressive Matrices*	-.01	.39*	.25	.42*	.39*	.52**	.40*	.42*	.56**
RAKIT	*Memory Span*	.22	-.01	.06	-.07	.03	.00	-.13	.05	.01

**. Correlation is significant at the 0.01 level (2-tailed),

*. Correlation is significant at the 0.05 level (2-tailed).

Note. The Performance Total includes all WISC-III^{NL} Perfomance subtests, except Picture Arrangement.

In Table 4.5 the correlations between grouped TagTiles scores and the grouped WISC scores are presented; the correlations were calculated with the Partial Least Squares method. Because only Vocabulary was included as part of the Verbal Comprehension factor, a total score was lacking and therefore Verbal Comprehension is omitted.

All TagTiles task clusters show a considerable correlation with WISC Perceptual Organisation and Processing speed. Tasks 6, 7 and 8 (complex skills) show the largest correlation (r=.7) with Perceptual Organisation. Tasks 2, 3 and 4 (memory skills) correlate moderately with Processing Speed (r=.6) and Perceptual Organisation (r=.5). TagTiles basic skills (Tasks 1 and 5) show a moderate correlation with both WISC factors as well. Again, these correlations suggest that very similar skills are addressed in the sets of tasks, though combined differently because no isolated correlations were found.

Table 4.5. Partial Least Squares correlation coefficients between grouped scores of TagTiles and WISC-III^{NL} (N=32). TagTiles Tasks 1 and 5 (Basic Skills), 2, 3, 4 (Memory skills) and 6, 7, 8 (Complex skills) are correlated with WISC factors.

Grouped TagTiles tasks	Basic	Memory	Complex
WISC factors			
Perceptual Organisation	.5	.5	.7
Processing Speed	.4	.6	.4

Finally, the total scores of TagTiles were correlated with WISC-III^{NL} performance total scores. A PLS response plot (Figure 4.3) was generated, visualizing the correlation between actual average TagTiles scores on Tasks 2-8 and PLS calculated scores based on one component. WISC-III^{NL} scores were used as input to predict the TagTiles scores with the model. The model shows a 45% fit of with the actual TagTiles data and a correlation of .67 with a prediction precision of +/- 30. When three components are assumed in the model, a correlation of .91 is found (fit 83 %, prediction precision +/- 17). Using three components can be justified, since the correlations between clustered tasks do suggest there are three underlying factors (see Table 4.5). However, the risk of over-fitting is real, given the number of participants (N=32) in the study.

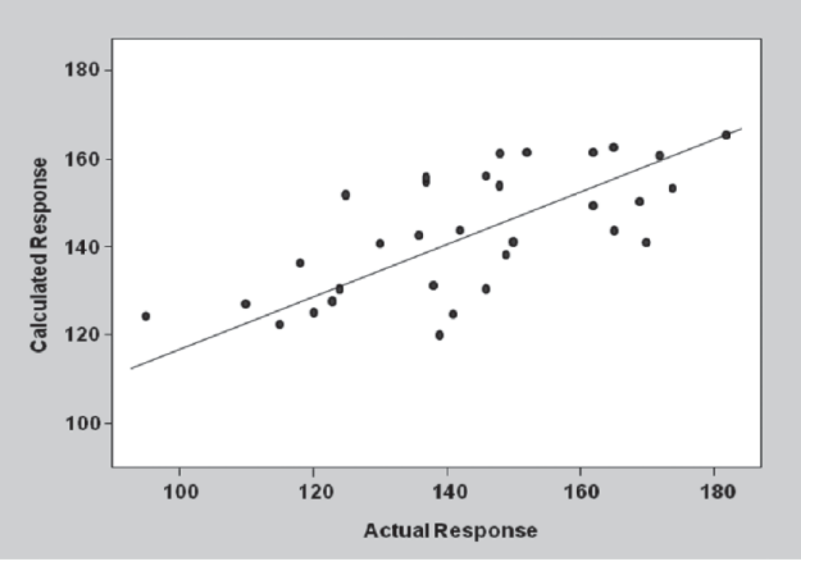


Figure 4.3. Partial Least Squares Response plot showing calculated responses of one component model correlated with actual responses of average scores for TagTiles Tasks 2-8.

Analysis of level complexity

To analyze the determinants of difficulty of the TagTiles patterns, each pattern was rated on a 3-point scale on different aspects (such as (a)symmetry) by two experimenters. The ratings were included in an ANOVA as factors, plus a factor 'participant' with 32 levels (N=32). The results show that 50% of variance in level completion times can be explained by the difficulty factors and 4% by participant properties, leaving the rest as unexplained variance. From the analysis we found that asymmetry in a pattern has the largest impact on difficulty (adding about 25 seconds to level completion time). A smaller but still strong impact has the space between cells, the number of colors used and mirror complexity, i.e. a diagonal mirror task is more complex than a horizontal mirror task. We could not apply a more thorough statistical analysis to investigate level complexity, since the impact on difficulty of different aspects of a pattern can only be estimated at this stage of the research.

4.4 Discussion

The findings of our study indicate that we can address similar skills with the TagTiles tasks as with the applied conventional psychological tests. This, even though there were no one-to-one correlations found between individual TagTiles tasks and the applied conventional psychological tests. The latter was expected, since the character of the tasks was not intended to be identical to conventional tests. The results can be seen as a confirmation of the suitability of the TagTiles tasks as a part of a psychometric instrument, though further validation is still needed.

Most of the hypothesized correlations were found in the results, though not all, and some unexpected correlations resulted from the study. We discuss specifically striking findings in more detail. For TagTiles, the results indicate that reproducing the sequence in reversed order (Task 4) was easier than in forward order (Task 3), shown by a significantly larger mean completion time and a smaller SD. This finding is difficult to explain, though one cause could be a learning effect that took place from Task 2 to Task 3, given the fact that they are both tasks that require sequential memory. Another explanation could be that the assignments in Task 3 were easier than the ones in Task 2, though when examining the applied patterns, they appear to be very similar.

Tasks 1 and 2 exhibited large correlations with Symbol Search, but for Task 5 and other tasks only a moderate correlation was found with this subtest, suggesting that selective attention, which is mainly needed for Symbol Search, is not as dominant in TagTiles Task 5 as we expected.

TagTiles correlations with Object Assembly performances were smaller than expected, and no significant correlations with Tasks 2, 3 and 7 were found. This contrasts our expectations and the findings with Block Design to some extent, but it confirms the expectation that the skills required for TagTiles are more similar to the skills required for Block Design than to those required for Object Assembly. A possible explanation for the finding that most of the TagTiles tasks exhibited significant correlations with Block Design (6 out of 8) would be that both have a time-component. Block Design scores include bonus points for fast completion, and TagTiles task scores are based on task completion time. However, in Object Assembly a scoring method similar to Block Design is applied, and with this sub test we found considerably fewer significant correlations (3 out of 8). Furthermore, Symbol Search does not have a time component and it has as many significant correlations with TagTiles tasks as Block Design (6 out of 8). This suggests that it is not mainly the time-component that causes TagTiles correlations with Block Design, but it may indicate that shared cognitive processes caused this result.

No significant correlations of TagTiles tasks were found with RAKIT Memory Span. This may be explained by the fact that the applied RAKIT test uses concrete figures, addressing verbal memory, and the memory tasks in TagTiles consist of abstract visual patterns which may be more difficult to label and thus to memorize.

Task 1 of TagTiles showed no significant correlation with Raven, though for other TagTiles tasks we found larger and significant correlations, suggesting that Task 1 does not address any skills related to deductive and spatial reasoning. This is in line with our expectation that Task 1 addresses basic cognitive processes (eye-hand coordination, motor skills) and reaction time. Given the finding that the average total TagTiles score shows a significant correlation (.55 at $p<0.1$) with Raven average performance, quite a strong overlap in the required skills can be assumed.

All but one of the calculated correlation coefficients between TagTiles and the conventional psychological tests with at least a moderate sized coefficient were positive. Surprisingly, TagTiles Task 8 showed a significant negative correlation with WISC's Digit Span Forward. Task 8 targets mainly spatial reasoning and Digit Span Forward measures verbal sequential memory (short term). The negative correlation indicates that strong Digit Span forward performance occurs with weak Task 8 performance and vice versa. In the past, some have observed a similar correlation and linked this to dyslexia (Rugel, 1974; Habib, 2000; Helland & Asbjørnsen, 2004). However, at the moment the general consensus is that there is not such a link (e.g., as discussed by Rosen, 2003).

The current study included a relatively small amount of participants (N=32). To validate the tasks as an assessment instrument, a study that includes a larger test sample is needed to be able to perform analyses with larger statistical power. This will also enable a better understanding of the properties that determine the difficulty of a TagTiles pattern-task combination. A better understanding would enable identification of those patterns that are most indicative of a child's developmental level for a certain set of skills.

The results of the current study are promising with respect to improvements on the testing practice issues that we described. The observations during the study and the pilot test interview results indicated that children enjoy completing the TagTiles tasks, which solves especially the issue of lack of motivation. When conventional psychological tests are embedded in an enjoyable game, the results are expected to give a more realistic impression of the abilities, as a potential lack of motivation or stress are in that case not likely to influence the outcome. In a parallel study (Verhaegh, Fontijn, Aarts & Resing, 2012) we have tested the use of the TagTiles tasks embedded in a game, which included fully automated instruction and scoring. This study also proved that the game was a positive experience for the children.

Test issues related to the test leaders are for the most part eliminated if the instructions for the tasks are automated, for instance by letting the TagTiles system itself provide them.

Previous research has shown that tests such as the WISC do not cover all academically relevant skills (Ardila, Pineda & Rosselli, 2000; Duncan et al., 2007). Executive functioning, working memory and motor skills are covered to some extend but they seem to deserve more attention in IQ assessments. With TagTiles we assume to facilitate the assessment of these skills more easily, due to the physical interface, which offers more degrees of freedom in tasks, combined with very precise registration of object manipulations on the board. We have already shown to some degree that this is possible with the current set of tasks, and other studies with TagTiles, e.g. on dynamic testing, confirm this (Resing, Steijn, Xenidou-Dervou, Stevenson, & Elliott, 2011; Henning, Verhaegh & Resing, 2011). In the parallel study we have found that children find the tangible TagTiles interface easy to use (Verhaegh et al., 2012). The latter supports our conclusion that TagTiles offers a suitable interface to children and decreases the chance of interference caused by the additional effort that the user interface may impose.

Based on the results of the current study, it cannot be determined conclusively which cognitive skills are addressed by which TagTiles task. One could argue that given the fact that the TagTiles performance score is time-based, the TagTiles performances might be determined predominantly by processing speed. However, if processing speed would cause the correlations found at test level one would not expect the very distinguished correlations found between some TagTiles tasks and some subtests, such as the correlation between Task 1 and Symbol Search. Furthermore, some sub tests correlate significantly with the average TagTiles task score while others do not correlate with any TagTiles task, for instance Picture Completion and Digit Span Backward. The pattern of correlations found rules out that a systematic factor is fully responsible for the results found. Also, a correlation of .57 was found between the average TagTiles task score and the WISC-III^{NL} Performance sub test total score. It is unlikely that such a large correlation with the WISC-performance score is due to a generic skill like processing speed only. Finally, the fact that the 'complex' TagTiles tasks exhibit a large correlation (r=.7) with perceptual organization, and a much smaller correlation with processing speed (r=.4), supports this as well.

Each cognitive task involves working memory and processing speed in one way or another. This may cause, in part, the correlation found between WISC and TagTiles tasks. The training effect found in (Verhaegh et al., 2012) could then be due to promoting the effective use of the existing working memory skills and processing speed. Further research is needed to separate these two from other, more specific cognitive skills like visual, spatial, verbal and reasoning skills.

For assessment we conclude that TagTiles can be used to test at least part of the cognitive skills that are addressed with the applied conventional psychological measures, given the significant correlations that were found. Concerning the use for training purposes, we conclude that at least some skills improve, given the improved performances of children over multiple training sessions (Verhaegh, et al., 2012). These skills likely include at least working memory and processing speed. Previous studies have already shown that training of relevant skills such as working memory can improve aspects of intellectual functioning, in particular executive functioning and efficient use of WM (Alloway & Alloway, 2009; Jaeggi, Buschkuehl, Jonides & Perrig, 2008). This opens the exciting prospect that by practicing with TagTiles the performance on the mentioned skills may be enhanced, or these skills may be more effectively used. This means that it would be useful to investigate whether, after further refinement and validation, the TagTiles tasks can be used for assessment and training of specific cognitive skills.

Acknowledgments

The authors thank Julia Henning for her assistance in data collection, Robert van Herk for supporting implementation of the TagTiles tasks and Bert Schriever for consulting in statistical data analysis. We thank Maddy Janse and Saskia van Dantzig for carefully reviewing previous versions of the paper. We are especially grateful to all children who participated in the experiments and their teachers.

5

Study III

In-game assessment and training of nonverbal cognitive skills using TagTiles

Verhaegh, J., Fontijn, W. F. J., Aarts, E.H.L. & Resing, W.C.M. (2012). In-game assessment and training of nonverbal cognitive skills using TagTiles. *Personal and Ubiquitous Computing*. Special Issue on Child Computer Interaction; Epub ahead of print. Springer. DOI 10.1007/s00779-012-0527-0.

Abstract

We present a field study with a game for children called 'Tap the little hedgehog' which is played on the TagTiles console, a tangible electronic interface. The game was developed to train and assess cognitive skills and includes tasks which, in isolation, exhibit a high correlation with a number of subtests from the Wechsler Intelligence Scale for Children (WISC-III^{NL}). The tasks address a range of nonverbal skills by requiring children to perform different operations on abstract patterns such as copying, reproducing sequences from memory and mirroring patterns. In the current study we tested whether these tasks kept their ability to address these skills if included in a gaming context, whether children are able to play the game independently and whether they are motivated to play the game. The results of the study support the hypothesis that nonverbal IQ-scores, as measured by the Wechsler Intelligence Scale for Children, can improve by training with a game. Hence, games like 'Tap the little hedgehog' can be used to train specific skills and serve as a screening tool for these skills. The results also confirm that children can play the game independently and that they enjoy it. We further found that children quickly learn how to play the game and use the interface.

5.1 Introduction

5.1.1 Background

One aspect of Ubiquitous Computing (Weiser, 1991) is the unobtrusiveness of the computing technology, as is also emphasized by the vision of Ambient Intelligence (Aarts & Marzano, 2003). In a previous article we introduced the concept of an integrated, fully unobtrusive system to support children in their development: the 'development support bubble for children' (Verhaegh, Fontijn, Aarts, Boer, & Van de Wouw, 2011). In this article we describe a gaming application that can be part of such an integrated system. The main property of the application is that while the game assesses and trains their cognitive skills, the children themselves just experience a fun gaming activity.

Current education is moving towards more independent forms of learning and there is an increasing need for tools to support personalized forms of teaching. Teachers typically do not have much time available to provide extra training to individual pupils. Parents prefer to offer their children toys or games that keep them active and engaged, and ideally teach them something on the side. Children themselves are most engaged by activities that they enjoy (Verhaegh et al., 2011). Taking the needs of all these stakeholders into account, tools that support independent, self-controlled and self-motivating learning and that are adaptive to individual needs, are very welcome. PC-based applications such as ‘Schoolpakket’ (http://www.ambrasoft.nl/school/), to be used as part of the school curriculum, or CogMed (Klingberg, Fernell, Olesen, Johnson, Gustafsson, Dahlström, Gillberg, Forssberg & Westerberg, 2005) as a remedial training to enhance working memory skills, already partially fulfill these needs. However, these solutions do not provide an interface that is easy to use for each pupil. Research has shown that for young children (age 7 or younger) controlling a mouse can be problematic and constitutes an additional challenge to the educational task that is offered (e.g. Verhaegh, Resing, Jacobs & Fontijn, 2009; Manches, O’Malley, & Benford, 2009; Hourcade, 2008). The result is that part of the training consists of learning to use the mouse, which takes precious attention away from the skill the training is supposed to develop (e.g. math). Removing this additional challenge is related to the notion of ‘offload cognition’: freeing up cognitive resources to help process domain relevant information, as described by Manches and O’Malley (Manches & O’Malley, 2012). Leveraging the developments in sensing technologies, surface and tabletop solutions are emerging that offer interfaces that are easier to control by children. They can offer direct access to a task, without requiring a spatial translation from mouse or keyboard input to screen output. Examples include the Smart Table (www.smarttech.com/table) and the Microsoft Surface (www.microsoft.com/surface). Another such product is the TagTiles console (www.serioustoys.com), which is controlled by manipulating physical objects on the TagTiles surface, bringing many benefits in the context of learning. For example, children can learn about spatial relations between objects ‘in the real world’. TagTiles is a console (see Figure 5.1) that includes a tabletop sensing board with an array of LED lights underneath and audio output. Tagged physical objects are used to provide system input. Any object can be used, depending on the requirements for the game, such as miniature boats, cars and puppets, and also more abstract shapes such as cubes and circles. New objects can be easily added. Removable plastic surface covers are used to display different game layouts on top of the TagTiles board.


Figure 5.1. TagTiles. The left image shows Ed’s game (training verbalizations of spatial relations); the right picture shows a child playing ‘Tap the little hedgehog’ (described in the remainder of the article).

The system provides immediate feedback on children's actions. For example, with the application 'Keer-op-Keer' children train multiplications and the surface of TagTiles shows how the outcome of the multiplication is reached in case children gave a wrong answer (or none at all), by lighting up the amount of squares that form the right answer. The game can be played with up to four players and the competitive aspect of the game leads to high engagement of the players, ensuring that they are all involved in training multiplications (Lathouwers, 2010).

To test electronic games for learning that employ physical learning materials, TagTiles is a very suitable tool. The system automatically keeps track of object manipulations on the board, which enables recording children's progress throughout the tasks. A detailed report with time logs can be obtained afterwards, that can be used for data analysis. Such a report contains quantitative data, i.e., it includes the order and timing of actions that were taken by a child to reach a solution (or not) in a task, providing rich information on a child's performance. This can be valuable information when studying the way children learn and it is a relatively easy way to collect data for pre-processing and analysis compared to, for example, visual recordings. Less detailed reports can also be obtained, for example providing a list of completed tasks, which could be useful feedback for a child's teacher.

5.1.2 Developing electronic learning applications with a physical interface

Our research has focused on the development of applications with a physical electronic (or tangible) user interface for learning and assessment. The work consists of three research phases, and each phase includes an empirical study. In the first phase we compared the use of a tangible user interface with a virtual interface in a puzzle task. This comparison showed that for young children physical objects offer a better interface for such a puzzle task than a PC mouse and a screen display (Verhaegh et al., 2009). In the second phase we investigated the effect of using pattern tasks (explained in the following section) on the TagTiles console to address several nonverbal cognitive skills, such as (working) memory and spatial reasoning (Chapter 4). This study included children aged 8-10 years and consisted of a pilot study (N=10) and an experiment (N=32). We found significant correlations between the performances on the TagTiles task and the performance of children on subtests of the nonverbal part of the Wechsler Intelligence Scale for Children III-NL (WISC-III^{NL} (Wechsler, 2005), namely for Block Design, Symbol Search and Coding. Some tasks also showed significant correlations with Raven's Progressive Matrices (Raven, Raven & Court, 2004) which is an intelligence test measuring deductive reasoning skills. The results of this study suggest that the tasks we developed can be used to train skills that are measured in IQ tests. In a parallel study, it was investigated how a scaffolding structure (a support structure) could be implemented into one of the TagTiles tasks that we had developed. An experiment was carried out with 15 children, demonstrating that at least part of the participants performed better with helpful hints that were provided by the TagTiles system (Henning, Verhaegh & Resing, 2011) than without these hints. In the third phase we examined whether the pattern tasks tested in phase two still provide a valid way to address the learning of specific skills when they become part of a game. We also investigated whether this game is experienced by children to be fun and intrinsically motivating. In the current article we present the results of the empirical study carried out in phase three.

5.2 Description of tangible educational game 'Tap the little hedgehog'

'Tap the little hedgehog', 'Tap de kleine egel' in Dutch, is a single player game developed to include the pattern tasks that we tested in the second research phase. In each of these tasks, a pattern of colored squares is presented on the TagTiles board. Different assignments are given to the player, e.g. to reproduce the pattern after it has disappeared, to reproduce a pattern in a specific sequence, or to mirror a pattern that is displayed on one side of the board onto the other side to create a symmetrical pattern. In Figure 5.2 all task names and an example are presented. The tasks were designed to address fine motor skills, eye-hand coordination, perception, attention skills and spatial memory. The higher difficulty levels also require reasoning.

1. Tag the pattern
2. Reproduce the sequence
3. Reproduce the sequence backwards
4. Reproduce the pattern from memory
5. Tag the pattern excluding the blinking squares
6. Extrapolate the pattern in a logical way
7. Mirror task (horizontal)
8. Mirror task (diagonal)

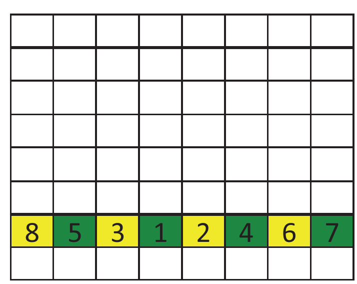


Figure 5.2. The names of the task assignments and an example pattern of Task 2 (level 2). In Task 2, the assignment is to reproduce a pattern which is displayed as a sequence of colored squares on the board. Before the player may start with the assignment, the pattern turns white. The numbers in the pattern are not actually displayed on the board.

We refer the reader to Chapter 4 for a complete description of all the tasks. Before the development and implementation of the game, we defined a number of requirements that we considered to be essential to create a fun and educational game that can be played by children independently. First, the eight pattern tasks described in Chapter 4 were taken as a given, as this study had shown the tasks' effectiveness in addressing a range of skills. However, in the current study, the difficulty of the task levels offered should adapt to the player's achievements. Second, we wanted to use a fantasy theme that encompasses the pattern tasks in a natural way, to minimize and simplify the instructions needed to understand the game play, to make the tasks more fun to play (Fontijn et al., 2007 and Malone & Lepper, 1987), and to create a natural 'flow' in the game. Third, we wanted to include a reward structure, to increase children's motivation to reach certain goals in the game. A final requirement to enable independent play was that the game application should include optional support that can be consulted by the child when needed, to repeat task instructions or to receive a helpful hint during a game.

5.2.1 Implementation of pattern tasks and game adaptivity

Eight different tasks were included in the game, employing all the tasks that were tested in a correlation study (Verhaegh, Fontijn & Resing, submitted). For the current study we added a number of patterns to the ones that had been used before, as we needed many more to enable enough variety for a game that can be played repeatedly, and to extend the range of difficulty. Previously the tasks included levels of increasing difficulty, which were offered in a fixed order independent of the player's performance. For the current study, the eight tasks were offered with three patterns of the same level of difficulty each (which we call 'task patterns' in the current article to avoid confusion). The eight tasks themselves also vary in their level of difficulty.

All tasks require the player to carry out an operation based on a pattern that consists of a number of squares. Each square contains a hedgehog (see Figure 5.3). The hedgehogs are part of the theme of the game. The patterns may include up to four colors which are displayed by light underneath the board surface (blue, red, green and yellow), corresponding to the colors of the playing pawns (small cubes) that are 'docked' in colored squares at the top of the cover. The pawns are used to provide system input by tagging the board. The players need to take pawns from the docking station at the start of the pattern assignment, and return them after completion to allow precise measurement of the pattern completion times.

The passing of time is represented by a lit up row of arrows that darken in sequence at the top of the cover. Each pattern should be completed within 60 seconds, otherwise the task pattern is aborted and the next one is started, or the next task is started if the aborted pattern was the third one for the task. After eight tasks with each three patterns, a new game round will start with task patterns of similar difficulty or of increased difficulty in case the player demonstrated he/she had mastered a skill level by reaching a certain score in the previous round. An example of the first pattern task is presented in Figure 5.4. The different tasks entail copying patterns (Task 1), reproducing them from memory (Task 2), reproducing sequences from memory (Tasks 3 and 4), copying patterns while ignoring the blinking squares (Task 5), extrapolating patterns in a logical way (Task 6) and mirroring patterns (Tasks 7 and 8); see Table 5.1 for the task names in the game and related skills.

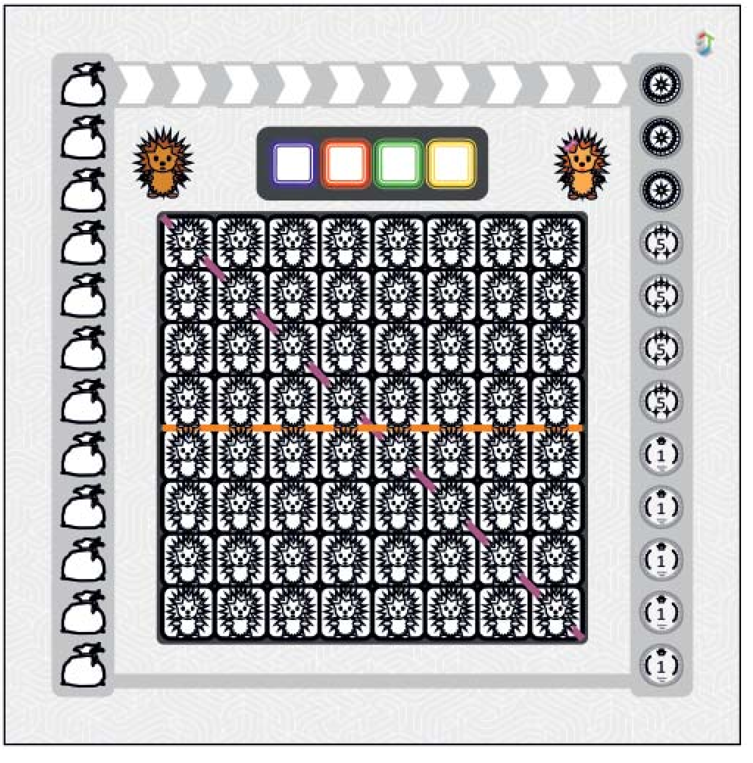


Figure 5.3. Surface cover of the game 'Tap the little hedgehog'.

In the correlation study we used one identical set of tasks and levels for all participants. In the current study we created a large number of task patterns and made the system randomly select from the pool of patterns for each participant. For each difficulty level, we used one pool of patterns. The design of the patterns was identical or close (i.e. rotated or in different colors) to the original patterns to ensure that they would require the same skills as much as possible. We assumed that rotating the pattern and exchanging colors would not significantly change the difficulty of the pattern or the addressed skills. This is an assumption to be tested, as will be discussed below.

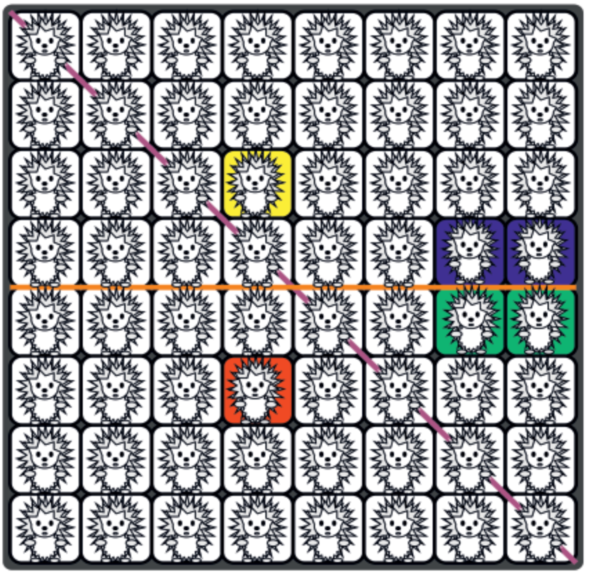


Figure 5.4. Pattern from Task 1: 'Feed the hedgehogs'. The (light) colors yellow, red, blue and green are displayed around the hedgehogs to indicate which color should be used to tag the squares with. The pattern on the board is displayed continuously throughout the completion of the pattern. The colored lines do not play a role in this task; they are used in the mirroring tasks.

Implementing game adaptivity

The game 'Tap the little hedgehog' implements the approach to balancing skill and difficulty as presented in Figure 2.5. For each task, three patterns of the same difficulty level are offered. The player earns points by completing a level successfully within 60 seconds. Faster completion yields extra points. After each task, the earned points are compared with a threshold value of points, to determine whether the level of difficulty should be increased in the next task to keep the child properly challenged. The points are different from the coins that a child earns (explained in section 5.2.4) and only used by the system. A task is eliminated in the next game round in case no points have been earned for this task in a prior round, to avoid frustration. What can be considered a form of scaffolding was provided through Tap and Tip (see section 5.2.3), functioning as tutors in the task.

5.2.2 Fantasy theme

The theme of the game is to save the hedgehogs by collecting as many coins as possible, displayed on the right side of the cover, to fill 'money bags' on the left side. The latter enable the player to 'buy' a sleeping place for the hedgehogs that protects them from the winter cold outside. The introduction to the game is provided by the system via a speaker. The theme adds to the game short term goals (earning coins) and long term goals (earning sleeping places) and the story of the theme supports the explanation of the tasks. Metaphors were used for tagging the squares in a pattern, such as 'feeding the hedgehogs', 'finding hedgehogs' sleeping place'; see Table 5.1 for all the task names. The tasks correspond to the ones presented in Figure 5.2, converted to the hedgehog theme. The choice of the theme was deemed suitable in terms of engagement for young children age 5 and up until about 10 years.

Table 5.1. The task names with related skills.

Task name	Skills
1. Feed the hedgehogs	Basic (perception, attention, motor)
2. Find back the disappeared hedgehogs 3. Feed the hedgehogs in the correct order 4. Feed the hedgehogs in reverse order	Memory (spatial, spatial sequential, and spatial sequential working memory)
5. Feed the hedgehogs that have not eaten yet	Basic (perception, attention, motor)
6. Extrapolate the line of hedgehogs 7. Find hedgehogs' sleeping place by straight mirroring 8. Find hedgehogs' sleeping place by diagonal mirroring	Complex skills (problem solving, spatial reasoning)

5.2.3 Support structure

The surface cover (Figure 5.3) shows two hedgehogs at the top of the cover, called 'Tap' (left) and 'Tip' (right). Before a task is played for the first time, Tap explains how the task should be played and a sample pattern is displayed. The instructions for a task are repeated if Tap is tagged after the start of a level. Tip can be tagged to receive visual hints during task completion (i.e. a part of the solution is shown). When there is no activity detected on the surface for 10 seconds, a prompt is given automatically ("Tag Tip if you would like to see the assignment again.").

5.2.4 Reward structure

After each successfully completed pattern, coins are awarded which light up on the right-hand side of the surface cover. The number of coins awarded depends on the speed at which the task was completed. If a pattern is completed within 60 seconds, applause is sounded and five coins are awarded. If a pattern is completed within 15 seconds, three bonus coins are awarded and if it is completed within 10 seconds five bonus coins are awarded. If the player collects 20 coins or more within the three levels of a task, 17 bonus coins are awarded. If all coins are lit, a money bag on the left-hand side lights up, indicating that enough is earned to buy a sleeping place for one hedgehog during winter time. To stimulate the player to complete the patterns as fast as possible, verbal praise is given to reward speed, e.g. "You played very well! You earned extra coins for speed."

5.3 Field test

A field test was set up to test the game. First, a brief pilot study with two participants aged four and five was carried out to fine tune the game and to ensure the children would be able to understand what was expected from them in the game. The right picture in Figure 5.1 depicts a child participating in the pilot test. The pilot test showed that young children could play the game for the largest part independently, providing an indication that the older children in the field test should be able to play it without help. Based on the observations in the pilot test, some changes were made to the timing of events during the task levels.

The main goal of the subsequent field test was to determine whether 'Tap the little hedgehog' can be used to address the same skills that were found to be addressed by the TagTiles tasks in the correlation study. We wanted to determine whether the potential to address these skills remained, given the added context, i.e. the game theme. The second goal was to determine whether children are intrinsically motivated to play the game, i.e. do they experience the game as enjoyable.

5.3.1 Participants

The participants were 52 children, 23 boys and 29 girls. The mean age of the girls was eight years and seven months (SD = 0.68) and the mean age of the boys was also eight years and seven months (SD = 0.75). The participating children were from three Dutch primary schools in grades two (26 children) and three (also 26 children).

The parents of the participating children had given permission for their child to participate in this study by signing an informed consent form. All participants passed a short version of Ishihara's test for color deficiency (Ishihara, 2005).

5.3.2 Experimental design and procedure

The field test consisted of four sessions, with approximately one week time between subsequent sessions. In the first session Ishihara's tests for color deficiency and a number of psychometric tests were administered. In the three subsequent sessions 'Tap the little hedgehog' was played for about 10 minutes, followed by administration of the evaluation questionnaire described below. In the fourth test session some of the children were also interviewed. The test sessions took place in a quiet area at the primary schools of the children.

5.3.3 Psychometric tests

A number of tests were applied to assess verbal and visual working memory abilities of the participating children, since these skills were expected to be relevant for the game tasks. We did not correlate these psychometric data with Tap performances, like we did in the correlation study, as the current test set up was not suited for this purpose. In the current test the participants were offered non-identical patterns, making comparison of the performances on these patterns difficult. Also, the offered patterns were on the whole easier than in the correlation study, resulting in less differentiation in the results.

We used part of the subtest Memory Span from the RAKIT (Revisie Amsterdamse Kinder Intelligentie Test (Bleichrodt, Drenth, Zaal & Resing, 1987)) to assess sequential visual memory for concrete figures. In this test a sequence of figures (e.g. tree, chair) is shown on paper and subsequently hidden from sight. The child is then asked to reproduce the sequence with cards displaying the figures. To assess visual search the Hidden Figures subtest of the RAKIT was administered, in which children have to find a hidden figure in a complex drawing. The subtest Digit Span from the WISC-III^{NL} (Wechsler, 2005) was used to test verbal memory skills. This test consists of two parts. In the first part (Digit Span Forward) a sequence of numbers is read to the child, and the task is to repeat the sequence. The length of the sequences increases with subsequent items of the test. In the second part (Digit Span Backward), the sequences that are read to the child are to be repeated in reversed order. The latter part measures verbal working memory skills.

5.3.4 The game

The children played the game without help from the experimenter. The experimenter was present during play though, to make sure that unexpected events (i.e. if a child would not start playing) could be observed and if needed, be responded to. The experimenter kept track of time, allowing children to play for about 10 minutes in each session. It depended on the speed of the child how many patterns would be completed. The tasks were always offered in the same order by the system. The task patterns, however, were not identical for the participants. These were taken automatically from a randomized pool of patterns, grouped per difficulty level. All task pattern completion times were registered by the system, to be used for later analysis.

5.3.5 Evaluation questionnaire

A questionnaire consisting of 16 items was applied to evaluate the children's gaming experience. The questionnaire is based on a selection of questions from the Intrinsic Motivation Questionnaire (Intrinsic Motivation Inventory, IMI) and the 'Fun questionnaire' (Stienstra & Hoonhout, 2002). The administration format was newly developed specifically for this test and the items were translated into Dutch. Three subscales were used: Interest/Enjoyment included 9 items (e.g. "I enjoyed playing the game"), Perceived Competence consisted of 4 items (e.g. "I am quite good at this game") and Effort and Importance included 3 items (e.g. "I put a lot of effort into this game"). The questionnaire was administered on the TagTiles board (see Figure 5.5 for the questionnaire surface cover). The questions were presented by the system via recorded speech, and the children used the pawns to indicate the answer of their choice. We only provided three answering options, since it is known that young children find it difficult to choose from many options. They tend to choose the extremes of a scale (Chambers & Johnston, 2002). The options were 'agree' and 'disagree', and an empty option in the middle. The latter was chosen because 'neutral' or 'don't know' were considered to be inappropriate labels. In addition we used colored icons (green, black and red): smileys, thumbs up/down, and a checkmark/cross to clarify the meaning of the labels. There were also two 'buttons', one to repeat the question and one to go to the next question.

Figure 5.5. Surface cover for the evaluation questionnaire.

5.4 Results of the field test

5.4.1 Game performance over three sessions

First we analyzed the game data. To compare the performances over the three sessions we used 710 seconds (11.8 minutes) as a cut-off point for the data of each session, because not all game-sessions had lasted exactly the same amount of time. Also the data from tasks 7 and 8 (mirroring tasks) were excluded from this analysis, due to the limited amount of data points collected for these tasks. Only few children had reached these tasks during the 10-minute game sessions. The results showed a significant main effect for mean pattern completion time, Wilks' $\lambda = .65$, $F(2, 32) = 8.57$, $p=.001$, $\eta p^2 = .349$ from game session 1 to 3 (see Figure 5.6), which means that the children become significantly faster over the three game sessions. Additional repeated contrast analyses of the mean time revealed that there was a significant difference between the first two sessions, $F(1, 33)= 4.71$, $p=.037$, $\eta p^2=.125$. There was also a significant difference between session 2 and session 3, $F(1, 33)= 10.22$, $p=.003$, $\eta p^2=.237$, showing that children become significantly faster with each subsequent session, completing more pattern tasks in the same amount of time. No difference between boys and girls was found in the pattern completion times (see Figure 5.6).

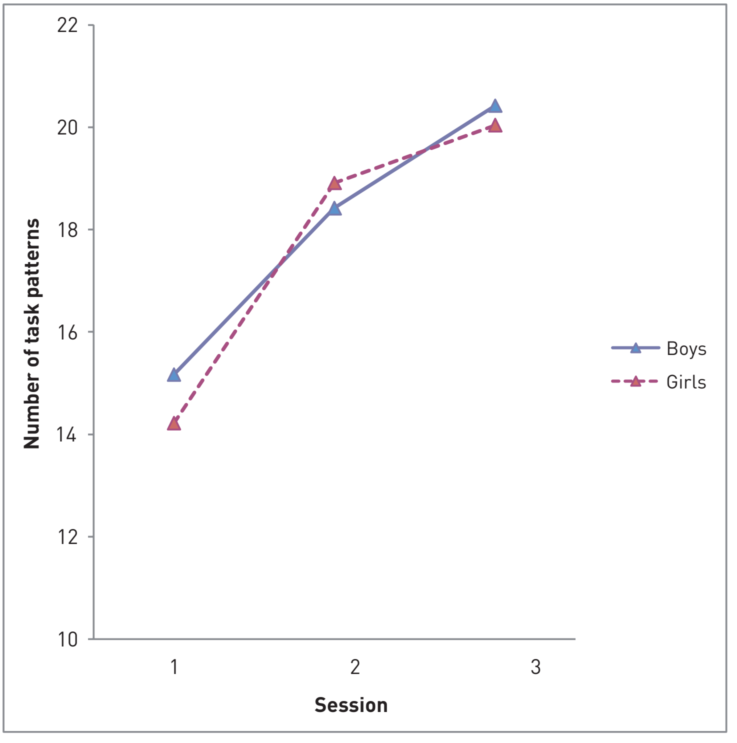


Figure 5.6. Mean number of task patterns completed per game session. Note that participants completed different task patterns per session, though of equal difficulty (matched per task).

5.4.2 Comparison of previous data set with current data set: linking results

In section 5.2.1 we mentioned the assumption that rotating the patterns and exchanging or rotating colors would not affect the level of difficulty of a pattern. Compared to the preceding correlation study, the Tap game includes many new patterns but also some of the original patterns were included. Hence, for specific patterns, we can make a direct comparison of the pattern completion times in the correlation study and those in the current study. The average pattern completion times derived from those in the preceding study are similar for most patterns (see Figure 5.7), which demonstrates that the inclusion of the tasks in the game does not affect the performance in itself. For some patterns the average level completion time was substantially higher than in the correlation study. This may be because in some special cases the added context does interfere, though it is not clear why. Another possible explanation is that in these cases applying the permutations did have an effect. For example, some patterns may be (or may be perceived to be) more difficult when offered in a horizontal orientation instead of a vertical orientation.

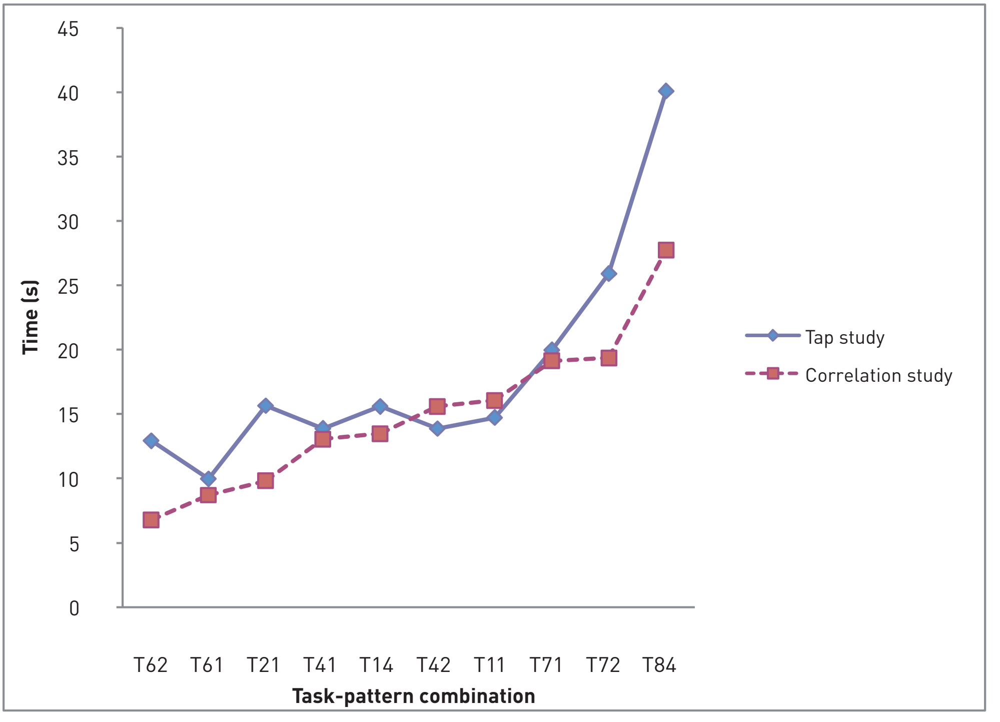


Figure 5.7. Average time per task sorted on average completion time in the correlation study, indicating relative difficulty. Txy refers to the task-pattern combination taken from the correlation study. E.g. T62 refers to Task 6 Pattern 2, which is matched with the data of all occurrences of the identical pattern in the current study (Tap).

5.4.3 Psychometric test results

In Table 5.2 the mean standardized scores and standard deviations for the 52 participants are displayed. The RAKIT standardized mean score per subtest is 15 and SD 5 (taken from the test manual (Bleichrodt et al., 1987)). For the WISC-III^{NL} subtests the mean score is 10 and SD 3 (taken from the test manual (Wechsler, 2005)). The psychometric subtests performances in this study match the average norm scores of the included age groups provided in the test manuals quite well.

Table 5.2. Mean scores and standard deviations (SD) per subtest.

Test	N	Mean	SD
RAKIT Hidden Figures (visual search)	52	13.38	6.04
RAKIT Memory Span (nonverbal working memory)	52	17.6	5.33
WISC-III^{NL} Digit Span (verbal working memory)	52	8.37	2.5

5.4.4 Evaluation results of the game experience: questionnaire

The questionnaires were filled out by part of the participants. The response rate per session is shown in Table 5.3. Note the decreasing standard deviation showing a decreasing data spread and thus more agreement in the answers. We initially analyzed the questionnaire as one scale. High ratings were found on the questionnaire, significantly increasing over the three sessions (a main effect for rating; Wilks' $\lambda = .569$, $F(2,12) = 4.546$, $p = .034$, $\eta p^2 = .431$). Contrast analyses of this main effect revealed that there was a non-significant difference between session 1 and session 2 ($F(1, 13) = .157$, $p = .699$, $\eta p^2 = .012$), but a significant difference between session 2 and 3 ($F(1, 13) = 7.76$, $p = .015$, $\eta p^2 = .374$). This indicates that the participants enjoyed playing Tap, and also that they seemed to enjoy it more when they played it more than once. Since we included items that were posed positively, but also reversed items (e.g. "I found the game childish to play with"), we could determine whether children answered consistently, which was the case. This suggests that children really listened to the questions before answering, and that they understood them.

Table 5.3. Intrinsic Motivation Questionnaire results. Mean total of the ratings (M) and standard deviation (SD); the maximum total was 16.

Session	N	Mean	SD
1	26	11.29	5.24
2	26	12.87	2.39
3	14	14.14	1.17

Split scores per subscale show that ratings on Interest/Enjoyment yielded the highest scores in each of the three sessions. Ratings on the Perceived Competence subscale show the largest increase over the sessions, 39% from the first to the third session. Remarkably, in the second session Effort/Importance was rated lower than in the first and the third session (see Figure 5.8).

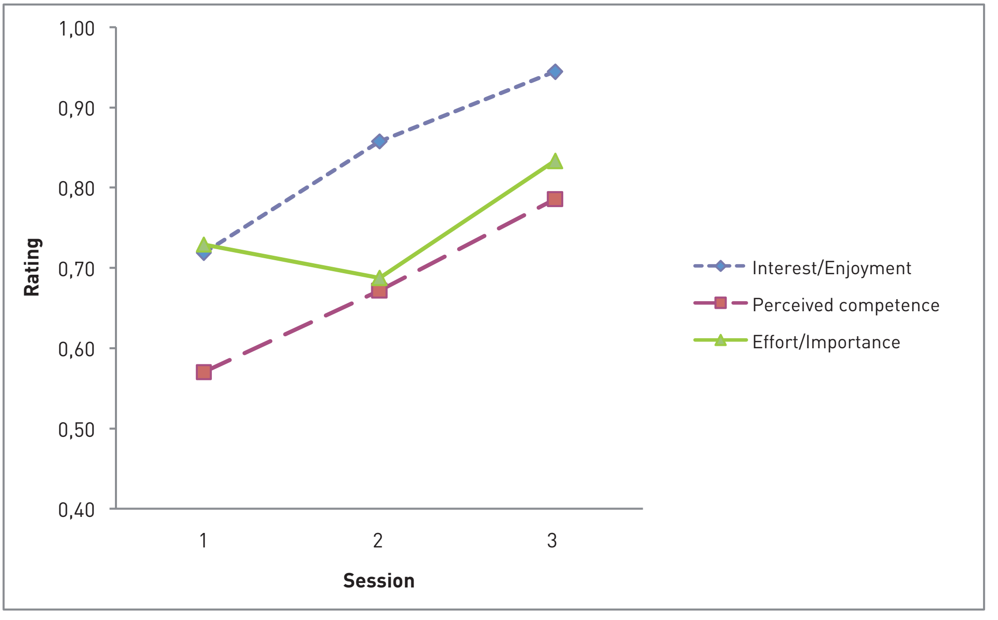


Figure 5.8. Questionnaire results. Normalized average ratings are depicted per session per subscale. It should be noted that the normalized ratings can take values from 0-1.

5.4.5 Evaluation results of the game experience: interview

Fourteen participants were interviewed after the final play session with 'Tap the little hedgehog'. We used a semi-structured interview asking children various questions about the game. To the question what they thought of the game, all of them said they regarded it to be fun to play. When asked what they considered to be the most fun, four answered 'working with the pawns', four liked everything and the rest mentioned various other things. Ten participants indicated that the tasks were too easy to play, but when asked whether they would like to see something changed in the game only two indicated that it should be made more difficult. The mirror tasks (Tasks 7 and 8) were often mentioned as the most difficult tasks. The story about the hedgehogs was appreciated by nine children (one said it was helpful), as confirmed by the questionnaire ratings, but the others deemed it boring, sad or not funny. The 'help' that could be received by tapping Tip or Tap was appreciated by 12 participants. The interview results are largely in line with the children's questionnaire responses.

5.5 Discussion and Conclusions

With the combined results of the correlation study (Chapter 4) and the current study we present evidence that TagTiles tasks applied in a game context can be used to assess and train a range of nonverbal skills. The observed similarity in the level completion times in the current and the correlation study suggests that the added game context does not, or at least not to a significant degree, change the gist of the tasks, indicating that they address the same skill set. This applies to most of the patterns that we could compare between both studies, even though we could not compare many patterns of higher difficulty level. In addition, the significant increase in speed over three sessions of play, suggests that Tap can be used as a means to train the skills that are addressed. From the study described in (Verhaegh et al.,2009) we know that the increase is not due to learning how to use the interface. TagTiles offers a direct interface to the tasks addressing the targeted skills.

The positive results from the questionnaires and interviews indicate that the children persistently enjoy the game, even after repeated play. Intrinsic motivation ratings from the questionnaire were high; and highest in the third session which may indicate that children who first did not like the game so much, start to enjoy it over time. The high ratings also suggest that our attempt to create a sense of accomplishment by including (sub) goals and rewards in the game, actually works.

During the game sessions no guidance was provided by the test leaders, demonstrating that children are able to complete the game tasks independently. This is also illustrated by the decreasing standard deviations from the mean pattern completion times after the first session, which suggests that the participants learned how to play the game after only one game session of 10 minutes.

Another TagTiles study that was conducted to evaluate scaffolding (Henning et al., 2011) showed that if needed TagTiles can provide hints to help the child to continue a task. The results of the interviews in this study indicate that this support is well appreciated.

The results of the correlation study demonstrated that TagTiles correlates significantly with WISC-III^{NL} performance. The current study showed that performances on Tap improve over subsequent play sessions, and that the tasks maintain their essence in the gaming context. These results support the assumption that nonverbal IQ-scores, as measured by the Wechsler Intelligence Scale for Children, can improve by training with Tap. This implies that applications like Tap can be used for example in a classroom setting to let children train skills independently.

The results also indicate that Tap can be used as a tool for screening of skills, to give a rough indication of the performance level. As such, Tap can be a tool in early detection of shortfalls in one of the addressed domains (i.e. spatial (working) memory, spatial reasoning, selective attention and perception), so timely measures can be taken. The usefulness of the TagTiles system in a specific type of assessment, dynamic testing, is already discussed (Resing, Steijn, Xenidou-Dervou, Stevenson & Elliott, 2011).

Considering the use of TagTiles as a means to train skills, the physicality of the interface seems a very powerful aspect in itself to support learning. This was also confirmed in a recent longitudinal neuro-imaging study where it was found that 'general verbal and non-verbal abilities are closely linked to the sensori-motor skills involved in learning' (Ramsden, Richardson, Josse, Thomas, Ellis, Shakeshaft, Seghier & Price, 2011). That study showed that non-verbal IQ changed with grey matter in a region that was activated by finger movements. When a higher grey matter density was found in this brain region, higher scores on the performance score of WISC were also found.

The version of Tap that we used in our study can be refined to become a tool that is better tuned to enable screening. This does require a more precise evaluation of the difficulty level of patterns. We can then use a few 'signal' patterns per task for assessment, and the remainder of the patterns used would not need a very fine-grained tuning or analysis, they would just serve for practice and fun. A follow-up study including signal patterns and a larger test sample is needed to further develop Tap in this direction.

Our studies so far have shown that a game like Tap can be valuable in supporting the skill development of children. We have also shown that Tap is a great example of a tool with an unobtrusive interface that gives direct access to the task of interest, making it especially suitable for the young users and as a part of a personalized development support bubble for children (Verhaegh et al., 2011).

Acknowledgements

The authors would like to thank Else Vink and others for their assistance in the data collection and Kirsten Touw for her support in the data analysis. We would like to thank also Privender Saini and Maddy Janse for their valuable comments on earlier versions of this paper. We are also grateful to the children who participated in the game tests.

6

General conclusions

In this dissertation the development and the use of TUIs for educational purposes has been described. First, it was investigated whether TUIs can be effective educational interfaces (Chapter 3). Then, we tested the hypothesis that specific (cognitive) skills can be addressed with such an interface (Chapter 4). Finally, it was investigated whether, based on the results of the studies in Chapters 3 and 4, a self-motivating and adaptive game could be made (Chapter 5). Together, the results of these studies support the hypothesis that we can create electronic games for a tangible user interface that are effective for developing cognitive skills of children. In the current chapter our findings are discussed as well as possible applications of the work in educational games and other areas. In addition, directions for future work are proposed.

6.1 General conclusions and discussion

The presented empirical studies are based on a number of theories regarding the cognitive development of children and learning. The relevance of hands-on learning (Montessori, 1912; Piaget, 1952; Goswami, 2008), a learning challenge that is continuously adapted to the child's developing skills (Vygotsky, 1978) and as a result, capitalizes on children's natural intrinsic motivation (Malone & Lepper, 1987; Wood et al., 1976) constitute the theoretical foundation of the empirical studies.

Prior research findings have not been consistent regarding the effects or possible benefits of tangible interfaces for learning (e.g, Marshall, 2007). Therefore in Chapter 3 it was investigated whether TUIs can be effective educational interfaces. This was tested with a study comparing the use of a PC with a virtual puzzle task, to a non-electronic tangible interface offering the same puzzle task. It was found that children (aged 5-7 years) were able to solve the tangible puzzle task on average almost twice as fast as the virtual task, and needed considerably less instruction for the tangible version. The results of the study support the hypothesis that tangible interfaces may offer a more suitable interface than PCs to educational tasks, at least for young children. Our results are in line with related work carried out by others (e.g., Freer Weiss, 2006; Khandelwal, 2006; Moyer, 2001; Olkun, 2003; O'Malley & Stanton-Fraser, 2004) supporting the positive effect of a tangible interface on educational tasks. The results further indicated that tasks involving the use of a PC mouse and a screen require more skills that are not relevant to the skills intended to be trained by the tasks. It is expected that educational tasks with a tangible interface provide more direct access to the challenge that is actually intended, leaving out the 'extra challenge' of the use of a mouse, and the spatial translation to a screen (Verhaegh et al., 2009).

However, if it is the intention to train skills that require the use of a PC, such an interface should be used. Therefore, in the development of an educational application, it seems advisable to first determine which skills need to be addressed, then consider the intended target group (e.g. younger or older children), and then choose the most suitable interface.

In Chapter 4 the hypothesis that specific cognitive skills can be addressed with a TUI was tested. A correlation study was carried out, in which children played a set of tasks on the TagTiles console and also completed a range of non-verbal subtests taken from conventional psychological tests including WISC-III^{NL}, Raven and RAKIT. Most of the hypothesized correlations were found in the results, though not all, and some unexpected correlations emerged in the study. We concluded that the skills that are addressed by the used TagTiles tasks cannot be distinguished exactly, though there must be some overlap with the applied psychological tests. Possibly some part of this overlap is due to shared working memory and processing speed components, which are part of most cognitive tasks. For assessment purposes we concluded that TagTiles can be used to test at least part of the cognitive skills that are addressed with the applied conventional psychological measures, given the significant correlations that were found. Studies by others have shown that training of relevant skills such as working memory can improve aspects of cognitive intellectual functioning, in particular executive functioning and efficient use of working memory (Alloway & Alloway, 2009; Jaeggi, Buschkuehl, Jonides & Perrig, 2008). This opens the exciting prospect that by practicing with TagTiles the performance on the mentioned skills may be enhanced, or these skills may be used more effectively. This means that it would be useful to investigate whether, after further refinement and validation, the TagTiles tasks can be used for assessment and training of specific cognitive skills.

In the study described in Chapter 5, the results from the prior two studies were used to develop the TagTiles game 'Tap the little hedgehog'. The eight pattern tasks tested in the correlation study were taken as a given and the difficulty of the task levels was made, to some extent, adaptive to the player's task performances. A fantasy theme was added to encompass the pattern tasks in a natural way, to minimize and simplify the instructions needed to understand the game play and to make the tasks more fun to play (Fontijn et al., 2007; Malone & Lepper, 1987), and to create a natural 'flow' in the game. A reward structure was applied, to increase children's motivation to reach certain goals in the game. Independent play was supported by including automated and optional support that can be consulted by the child when needed, to repeat task instructions or to receive a helpful hint during the game.

The results of the empirical study with this game indicated that the added game context had not changed the essence of the tasks, as the task performances were similar to those in the correlation study. Therefore, the results support the hypothesis that TagTiles tasks applied in a game context can be used to assess and train a range of nonverbal skills. The observed similarity in the level completion times in this study and the correlation study suggest that the added game context does not, or at least not to a significant degree, change the gist of the tasks, indicating that they address the same skill set. In addition, the significant increase in performance over three sessions of play, suggests that Tap can be used as a means to train the skills that are addressed.

Educational game design requirements

To guide the development of educational game content, the requirements that were presented in Chapter 1 and 2 have been implemented in the described empirical studies with the TagTiles console. The use of these requirements is briefly reflected upon in this section.

Adaptivity To facilitate learning, it is important that the challenge offered is tailored to the abilities of a child (Vygotksy, 1978). The adaptivity that was implemented in the game 'Tap the little hedgehog' was found to support the children sufficiently for them to play the game independently. A partial implementation of adaptivity according to the described approach in Chapter 2 (Figure 2.5) was enough to enable independent play. More refined tailoring of adaptivity is expected to provide more detailed information on the child's actual performance level of specific skills. Further refinements of the adaptivity can include more fine-grained skill-level detection. For example, in the current version of the game, in case a task had not been completed successfully in the previous round, the task would be left out in the next game round. If skill-detection is more precise, it should become possible to offer a task that is only within reach for a child with proper scaffolding (e.g., Granott, 2005). This is supported by the studies of Resing and Elliott (2011) and Henning et al., (2010) who developed an adaptive training in inductive reasoning, using scaffolding principles based on cognitive task analyses.

Self-motivation Self-motivation has been claimed to enhance learning (Deci, Koestner & Ryan, 1999; Lepper, Corpus & Iyengar, 2005). At the same time, research has shown children's self-motivation decreases over the years in the current school systems (Lepper, Corpus & Iyengar, 2005; Malone & Lepper, 1987). Games are usually experienced as fun and therefore learning tasks encompassed in a game can be useful tools for testing and training skills. The results of our studies have demonstrated that both the separate TagTiles tasks and the game 'Tap the little hedgehog' were engaging to the children. In the correlation study with the separate TagTiles tasks the willingness to keep playing continued in most cases until the children had reached the most difficult tasks. In case it became too difficult and because no help was offered, a number of children indicated that they no longer wanted to play. This was not the case anymore in the study with 'Tap the little hedgehog', suggesting that the inclusion of the tasks in a game including a theme and help options succeeds to keep children motivated to play independently. This was confirmed by the answers children provided in the motivation questionnaire and in the interview results.

Addressing the intended skills and involving domain experts Educational tasks are expected to be less effective in case the cognitive load is targeted inappropriately. It is known that cognitive processing capacity is limited (Sweller, van Merriënboer & Paas, 1998) and thus the intended skills need to be targeted as precise as possible to enable effective testing or training. Part of the cognitive load that a task elicits is due to the task interface and therefore the interface needs to suit the task purpose and avoid unintended cognitive load. The results of the study in Chapter 3, where the use of a tangible interface was compared with a virtual interface, support the hypothesis that tangible user interfaces can be more effective learning interfaces than PC-based interfaces. It was concluded that tangible interfaces provide more direct access to a task, whereas virtual, PC-based, tasks can introduce an unwanted extra challenge, such as learning to control a PC mouse. To determine how the intended skills can be addressed, domain experts such as (remedial) teachers can be involved in the task design process. In the development of 'Tap the little hedgehog' domain experts were involved. Whether this is a requirement for such games was not tested within the scope of this dissertation. In the development of other educational games for the TagTiles console however, the inclusion of domain experts has been a fruitful and valuable way of working. Studies with the resulting educational games have shown that these games are effective in delivering both the desired educational yield as well as self-motivation (Lathouwers, 2010; Ter Braak, 2011).

6.2 Applications of the work

Skill development and assessment

The results of this dissertation are in the first place applicable in educational and assessment practices. Scholars interested in the development of electronic educational tools may find our implementation of cognitive theories into the TagTiles prototype helpful in the design of new educational tools. The described adaptive use of tangible computing allows putting the child and its specific needs at the center while maintaining a focus on educational effects. Cognitive, social and fine motor skills can thus be addressed in an integrated manner (Verhaegh et al., 2011).

Furthermore, as shown with the TagTiles studies, tangible computing can be used to enhance the normal educational process, but also to signal and treat shortfalls in development. The TagTiles environment is flexible and allows for intrinsically motivating educational applications with which the child can work both independently and collaboratively. At the same time the objects used in tangible interfaces are familiar to the children and connect to their natural way of play, thus creating a natural transition from playing at home through playing in a school environment to focused learning at school.

In the hands of domain experts, electronic tangible systems like the TagTiles console are very powerful tools for the integral and personalized development of children in the areas of cognitive, fine motor and social skills for assessment, education and therapy. Different skills can be trained in an integrated manner. For example, an application that is mainly targeting motor skills can be made more interesting and motivating by including a cognitive challenge in the task (Li et al., 2008). Furthermore, as the exercises can be presented in the form of attractive games, the children are intrinsically motivated to use them while their performance can be monitored unobtrusively. In the case of TagTiles, the tasks can be designed by domain experts in terms of hardware, but also software by using the dedicated Software Development Kit "ESPranto" (Van Herk et al., 2009). This kit allows easy programming of applications, without requiring programming knowledge. Finally, as professionals in the field like teachers and occupational therapists have pointed out (e.g. Li et al., 2008), tools like TagTiles can be used by the children unsupervised and hence as easily at home as in a more formal setting. Electronic reports of unsupervised use can be obtained afterwards. The use of the TagTiles console illustrates that electronic tangible interfaces are easy to use, easy to learn to use and offer great opportunities for personalization and contextualization of the developmental process. Next to showing that these benefits are real it was found that these benefits in particular are of importance to both parents and teachers (Verhaegh et al., 2011) and that both also recognize that electronic learning aids based on physical computing, like the TagTiles console, deliver on these benefits. The ability to deliver this combined set of benefits in an integral manner may be unique to tangible electronics. Further, we have pointed out the opportunities that are envisioned in the domains of assessment and training of particular skills for children with specific needs. In spite of these opportunities, tangible electronics are not applied yet in the field of assessment and training (Wijshoff, 2010). The number of PC-based applications used in practice is increasing though. Examples are the trainings of working memory such as CogMed (www.cogmed.com; Klingberg et al., 2005) and JungleMemory (www.junglememory.com). As described before though, the PC interface may not always be optimal for targeting specific skills and it can be difficult to use for children.

Enhancing game experiences

The results of this dissertation can also be used for the development of electronic games that are purely designed for entertainment purposes, to enhance the game experience. Although different aspects are assumed to contribute to the experienced fun (e.g. Malone & Lepper, 1978; Fontijn & Hoonhout, 2008), offering an appropriate challenge to the user seems to be a crucial factor. Games that include a challenge that adapts to the user's progression can be more interesting and motivating than games that are static, or only offer coarse levels of increasing difficulty. If the skill level of the player can be monitored while playing, a tailored challenge can be offered continuously. Boredom and frustration are avoided this way and by keeping the player on his toes continuously, the educational yield is also maximized. To implement an adaptive challenge into an electronic game, the principles described with the ZPD in Figure 2.5 can be helpful.

In the case of computer based board games, which are emerging such as Jumbo IPawn® and GameChanger from Identity Games™ (both to be used with IPad), it can be interesting to create challenges that are appropriate for each player individually. For example, in case a child would like to play with a parent the game can be enjoyable for both by steering the challenge such that it is engaging for both players, even if their skill levels are different (Kraaijenbrink, van Gils, Cheng, van Herk & van den Hoven, 2009).

Controlling complex devices

The results from this dissertation (Chapter 3) also indicate the usefulness of tangible interaction in controlling complex devices. Tangible interaction can support easy, playful ways of learning to use complex devices such as remote controls of TVs or complex lighting systems (Fontijn, 2007), making the use of manuals no longer necessary. Implicit control, where the device learns to anticipate the user's wishes, but also more natural explicit control of a device, can be facilitated with tangible interaction. A learning curve may be facilitated where the user first learns how to use the basic functionalities of a device, which, once mastered, are supplemented with more complex functionalities. This requires a device to detect the user's level of understanding, based on the way it is used. For example for elderly users of complex devices, initially reducing functionality may be helpful. The concept of the Zone of Proximal Development (Vygotsky, 1978) may be useful for developing this kind of applications as well. Scaffolding can help the user to become more skilled at controlling the device, and eventually enable the use of complex functionalities independently.

6.3 Future research directions

Over the past decades the demand for early testing of children's development has increased. This demand has led to a debate questioning whether these tests can be done appropriately and whether the results will lead to educational benefits (e.g. Dietel, 2003). Early signaling of shortfalls in children's cognitive development is an actual issue and appropriate assessment instruments to facilitate this are needed. As discussed in this dissertation, there are several reasons why conventional tests may not be optimal for young children, such as stress because of the testing situation, or a lack of motivation to participate or do well on a test. Irregularities in the test administration and scoring can occur as well and not all relevant skills can be assessed properly with existing tools (e.g. motor skills) and also the test format (e.g. pen/paper, PC) may offer an added challenge that varies per child. Psychological tools that are embedded in tangible electronic games offer a non-threatening, playful solution for the assessment of young children.

A game like 'Tap the little hedgehog' is suitable for offering a test of multiple cognitive skills, and it can be tuned to the needs of the child. The version of 'Tap the little hedgehog' that we used in our study can be refined to become a tool that is better tuned to enable screening of cognitive skills.

To enable the use of the game in the testing practice, options to further develop and validate the game should be explored. Part of this can be the validation of the visual patterns like the ones used in our studies, enabling determination of the skill level of children in specific cognitive tasks. The patterns can be systematically varied to investigate the effects this has on children's performances. This requires the use of many different patterns with a larger test sample. A resulting set of 'signal patterns' can be used in any game theme deemed suitable for the intended target group. Once this set of signal patterns is found, these may again be correlated with specific psychological tests addressing specific cognitive processes, to identify which skills can be addressed exactly besides more general skills such as working memory and processing speed. For example, Mammarella and colleagues (2008) have found evidence for multiple components of working memory that may exist using a similar approach. This way diagnostic use of the tool can be made more specific.

For the development of skills it was concluded that it is possible to train skills by using a game like 'Tap the little hedgehog'. With the current version, it is not clear yet which skills can be trained exactly, but it may be used as a general tool for working memory and processing skills development. Once further validation of the signal patterns has taken place, it can also be used more deliberately for training purposes of more specific cognitive skills. Aspects related to the design of the game (hardware and software), such as the use of themes may also be further explored, including further testing with children. Ideally, this leads to applications that are fully tailored to children's personal interests and individual learning needs.

As shown, when designing the educational tools and methods of the future, putting the child and its natural way of developing at the center offers great benefits. The child will be more motivated and at the same time the educational yield will be higher and more targeted. Electronic tangible systems like the TagTiles console offer integral, personalized development of children in the areas of cognitive, fine motor and social skills for assessment, education and therapy, in a manner that builds on natural forms of play of children.

Bibliography

Aarts, E.H.L. & Marzano, S. (eds). (2003). *The New Everyday: Visions of Ambient Intelligence*, 010 Publishing: Rotterdam, The Netherlands.

Abt, C.C.. (1987). *Serious Games.* University Press of America: Lanham, MD.

Africano, D., Berg, S., Lindbergh, K., Lundholm, P., Nilbrink, F., & Persson, A. (2004). Designing tangible interfaces for children's collaboration. In *proc. of CHI '04 Extended Abstracts on Human Factors in Computing Systems, ACM Press,* 853-868.

Alloway, T. & Alloway, R. (2009). The efficacy of working memory training in improving crystallized intelligence. Available from Nature Precedings http://hdl.handle.net/10101/npre.2009.3697.1.

Alloway, T.P. & Alloway, R.G.(2010). Investigating the predictive roles of working memory and IQ in academic attainment, *Journal of Experimental Child Psychology, 106* (1), 20-29.

AmbraSoft. retrieved October 15, 2011, from http://www.ambrasoft.nl/school/.

Antle, A. N. (2007). Designing tangibles for children: what designers need to know. In *proc. of CHI '07 Extended Abstracts on Human Factors in Computing Systems*. ACM, New York, 2243-2248.

Appl, D.J. (2000). Clarifying the Preschool Assessment Process: Traditional Practices and Alternative Approaches. *Early Childhood Education Journal, 27* (4), 219-225.

Ardila, A., Pineda, D. & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology, 15* (1), 31-36.

Bayliss, D. M., Jarrold, C., Gunn, D. M., & Baddeley, A. D. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology-General, 132* (1), 71-92.

Berk, L. (2008). *Infants and children (sixth edition*). Essex, UK: Pearson Education.

Biddell, T.R., & Fischer, K.W. (1992). Beyond the stage debate: Action, structure, and variability in Piagetian theory and research. In R.J. Sternberg & C.A. Berg (Eds.), *Intellectual development* (pp. 100-140). Cambridge, UK: Cambridge University Press.

Bleichrodt, N., Drenth, P.J.D., Zaal, J.N. & W.C.M. Resing (1984). RAKIT (Revisie Amsterdamse Kinder Intelligentie Test). Lisse: Swets & Zeitlinger.

Bruner, J. S., & Olver, R. (1963). Development of equivalence transformations in children. In H.F. Wright & J. Kagan (Eds.), *Basic cognitive processes in children. Monographs of the Society for Research on Child Development, 28,* 125-142 (Series 86).

Bruner, J.S. (1964). The course of cognitive growth. *American Psychologist, 19*, 1-15.

Bruner, J.S. (1973). *Beyond the information given: Studies in the psychology of knowing.* New York: Norton & Company.

Chambers, C.T. & Johnston, C. (2002). Developmental differences in children's use of rating scales. *J. Pediatr. Psychol., 27* (1), 27-36.

Cogmed Working Memory Training. Retrieved September 30, 2012, from http://www.cogmed.com.

Cowan, N., Saults, J. S., & Elliott, E. M. (2002). The search for what is fundamental in the development of working memory. In *Advances in Child Development and Behavior, 29,* 1-49.

Csikszentmihalyi, Mihaly (1990). *Flow: The Psychology of Optimal Experience*. New York: Harper and Row.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125* (6), 627-668.

Deci, E. L., Koestner, R., & Ryan, R. M. (2001). Extrinsic rewards and intrinsic motivation in education: reconsidered once again. *Review of Educational Research, 71* (1), 1–27.

Dietel, R. (2003). Issues in testing very young children. Center for the Assessment and Evaluation of Student Learning, *Assessment Brief, Number 5,* January 2004. Retrieved from http://www.caesl.org/briefs/Brief5.pdf.

Dienes, Z. (1964). *Building up mathematics* (2nd ed.). London: Hutchinson Educational.

Dijksma, F. T. R., Resing, W. C. M., Roig, I., Van Herk, R., Meijles, J. & Verhaegh, J. (2008). *Electronic dynamic testing with a puppet series completion task*. Poster presented at the International School Psychologist Association Conference, Utrecht, 8-12 Juli 2008.

Dourish, P. (2004). What We Talk About When We Talk About Context. *Personal and Ubiquitous Computing, 8* (1), 19-30.

Driscoll, M.P. (1994). *Psychology of learning for instruction.* Needham Heights: Allyn & Bacon.

Duckworth, A. L., Quinn, P. D., Lynam, D. R., Loeber, R., & Stouthamer-Loeber, M. (2011). Role of test motivation in intelligence testing. In *Proceedings of the National Academy of Sciences*, *108* (19), 7716-20.

Duncan, G.J. Dowsett, C.J. Claessens, A. Magnuson, K. Huston, A.C. Klebanov, P. Pagani, L.S. Feinstein, L. Engel, & M. Brooks-Gunn, J. (2007). School Readiness and Later Achievement. *Developmental Psychology, 43* (6): 1428-1446.

Ekström Ahl, L., Johansson, E., Granat, T. & Borgren Carlberg E. (2005). Functional therapy for children with Cerebral Palsy: an ecological approach. *Developmental Medicine & Child Neurology, 47,* 613-619.

Elliott J.G., Grigorenko E.L. & Resing, W.C.M. (2010). Dynamic Assessment. In: Penelope Peterson, Eva Baker, Barry McGaw, (Eds), *International Encyclopedia of Education, 3,* pp. 220-225. Oxford: Elsevier.

Evers, A., Lucassen, W., Meijer, R. & Sijtsma, K. (2009). *COTAN beoordelingssysteem voor de kwaliteit van tests (geheel herziene versie).* Amsterdam: NIP.

Farnham-Diggory, S. (1992). *Cognitive Processes in Education, 2nd Edition*. New York, NY: HarperCollins Publishers, Inc.

Field, A. (2009). *Discovering statistics using Spss.* London: Sage.

Fishkin, K.P. (2004). A taxonomy for and analysis of tangible interfaces. *Personal Ubiquitous Comput. 8,* 5, 347–358.

Fontijn, W. F. J., & Mendels, P. (2005). StoryToy the interactive Storytelling Toy. PerGames workshop. In *proceedings of International Conference on Pervasive Computing,* Munich, Germany, 11 May 2005, pp. 37-42.

Fontijn, W.F.J. & Hoonhout, H.C.M. (2007). Functional Fun with Tangible User Interfaces. In *Proc. of First IEEE International Workshop on Digital Game and Intelligent Toy Enhanced Learning, DIGITEL '07*, Jhongli City, Taiwan, p. 119-123.

Fontijn, W.F.J. (2007). Ambient Intelligence. In Bemelmans, T.M.A., van Keulen, M., Kusters, R.J. and Looijen, M., (Eds.). *ICT-Zakboek*. PBNA Polyzakboekjes (pp. 367-377). Den Haag: Reed Business Information.

Fontijn, W.F.J. & Verhaegh, J. (2008). Patent Application WO2010033032, *Game device and method of assessing skill of a user*, Serious Toys B.V., 22-09-2009, retrieved from http://www.wipo.int/patentscope/search/en/WO2010033032.

Freer Weis, D. (2006). The rationale for using manipulatives in the middles grades. *Mathematics teaching in the middle school, 11*, 238-242.

Fröbel, F. (1826*). On the education of man (Die Menschenerziehung).* Keilhau/Leipzig: Wienbrach.

Gallahue, D. L., & Ozmun, J. C. (2002). *Understanding motor development: Infants, children, adolescents, adults* (5th ed.). Boston: McGraw-Hill.

Geirland, J. (1996). "Go With The Flow". Wired magazine, September, Issue 4.09., retrieved from http:/ /www.wired.com/wired/archive/4.09/czik_pr.html).

Gleitman, H., Fridlund, A.J. & Reisberg, D. (1999). *Psychology* (5th edition). London: W.W. Norton & Co Ltd.

Goswami, U. (2004). Annual review Neuroscience and education. *British Journal of Educational Psychology, 74,* 1–14.

Goswami, U. (2008). *Cognitive development: The learning brain.* Hove, UK, Psychology Press.

Granott, N. (2005). Scaffolding dynamically towards change. Previous and new perspectives. *New Ideas in Psychology, 23,* 140–151.

Grissmer, D., Grimm, K.J., Aiyer, S.M., Murrah W.M. & Steele, J.S. (2010).Fine Motor Skills and Early Comprehension of the World: Two New School Readiness Indicators, *Developmental Psychology, 46 (5),* 1008-1017.

Habib, M. (2000). The neurological basis of developmental dyslexia: An overview and working hypothesis. *Brain, 123* (12), 2373-2399.

Helland & Asbjørnsen (2004). Digit span in dyslexia: Variations according to language comprehension and mathematics skills. *Journal of Clinical and Experimental Neuropsychology, Vol. 26* (1), 31–43.

Hembree, R. (1988). Correlates, Causes, Effects, and Treatment of Test Anxiety, *Review of Educational Research, 58 (1),* 47-77.

Hendrix, K., van Herk, R., Verhaegh, J., & Markopoulos, P. (2009). Increasing children's social competence through games, an exploratory study. In *Proceedings of the 8th international Conference on interaction Design and Children* (Como, Italy, June 03 - 05, 2009). IDC '09. ACM, New York, NY, 182-185.

Hengeveld, B.J. (2011). *Designing LinguaBytes : a tangible language learning system for non- or hardly speaking toddlers.* (Doctoral dissertation). Available from TU/e repository; http://repository.tue.nl/715535, Technische Universiteit Eindhoven.

Henning, J.R., Verhaegh, J. & Resing, W.C.M. (2011). Creating an individualized learning situation using scaffolding in a tangible electronic series completion task. *Educational and Child Psychology, 28* (2), 85-100.

Hourcade, J.P. (2008). Interaction Design and Children. *Foundations and Trends in Human–Computer Interaction, 1* (4), 277-392.

Intrinsic Motivation Inventory. Retrieved July 1, 2012 from
http://www.selfdeterminationtheory.org/questionnaires/10-questionnaires/50.

Ishihara, S. (2005). *The series of plates designed as a test for colour deficiency*, Concise edition, Kanehara Trading Inc.: Tokyo, Japan.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. In *Proc. of the National Academy of Sciences of the United States of America, 105 (19),* 6829–6833.

JungleMemory. Retrieved July 21st 2011 from http://junglememory.com/.

Khandelwal, M. (2006). *Teaching table: a tangible mentor for pre-kindergarten math education.* Atlanta, Georgia: Georgia Tech University: Master Thesis.

Khandelwal, M. & Mazalek, A. (2007) Teaching Table: A tangible mentor for pre-K math education. In *Proc. of the 1st international Conference on Tangible and Embedded interaction*, ACM Press (2007), 191-194.

Klingberg, T., Fernell E., Olesen, P.J., Johnson, M., Gustafsson, P., Dahlström, K., Gillberg, C.G., Forssberg, H. & Westerberg, H. (2005). Computerized Training of Working Memory in Children With ADHD—A Randomized, Controlled Trial. *J. Am. Acad. Child Adolesc. Psychiatry, 44* (2), 177–186.

Kraaijenbrink, E.B., Gils, F.M.D.M. van, Cheng, Q., Herk, R. van & Hoven, E.A.W.H. van den (2009). Balancing skills to optimize fun in interactive board games. In *Proc. of Human-Computer Interaction (Interact 2009)*, Lecture Notes in Computer Science, 5726, pp. 301-313. Berlin: Springer.

Kuentzel, J.G., Hetterscheidt, L.A. & Barnett, D. (2011). Testing intelligently includes double-checking Wechsler IQ scores, *Journal of Psychoeducational Assessment, 29* (1), 39-46.

Lathouwers, E. (2010). *Spel in de rekenles.* Internal thesis report, Serious Toys. B.V.

Lehto, J.E., Juujarvi, P., Kooistra, L. & Pulkkinen, L. (2003). Dimensions of executive functioning: evidence from children. *British Journal of Developmental Psychology, 21* (1), 59-80.

Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and extrinsic motivation orientations in the classroom: Age differences and academic correlates. *Journal of Educational Psychology, 97*, 184–196.

Levine, S. C., Huttenlocher, J., Taylor, Al., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology, 35*, 940-949.

Li, Y., Fontijn, W.F.J. & Markopoulos, P. (2008). A tangible tabletop game supporting therapy of children with Cerebral Palsy. In *Proc. of the 2nd International Conference on Fun and Games 2008,* p. 182-193.

Malone, T. W., & Lepper, M. R. (1987). Making Learning Fun: A Taxonomy of Intrinsic Motivations for Learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitute, Learning and Instruction: III. Conative and affective process analyses* (pp. 223-253). Hillsdale, NJ: Erlbaum.

Malone, T.W. (1980). What makes things fun to learn? Heuristics for designing instructional computer games. In *Proc. of the 3rd ACM SIGSMALL symposium (SIGSMALL '80)*. ACM, New York, NY, 162-169.

Malone, T. W. (1982). Heuristics for designing enjoyable user interfaces: Lessons from computer games, *Conference on Human Factors in Computing Systems*, pp. 63-68.

Mammarella I.C., Pazzaglia F., & Cornoldi C. (2008). Evidence for different components in children's visuospatial working memory. *British Journal of Developmental Psychology, 26,* pp. 337-355.

Manches, A., O'Malley, C., & Benford, S. (2009). The role of physical representations in solving number problems: a comparison of young children's use of physical and virtual materials. *Computers & Education, 54* (3), 622-640.

Manches, A. & O'Malley, C. (2012). Tangibles for learning: A representational analysis of physical manipulation. *Personal and Ubiquitous Computing, 16*, 405-419.

Marshall, P. (2007). Do tangible interfaces enhance learning? In *Proceedings of the 1st International Conference on Tangible and Embedded interaction* (pp. 163-170). New York: ACM Press.

Meijer, J. (1996). *Learning potential and the fear of failure. A study into the predictive validity of learning potential and the role of anxious tendency.* Unpublished doctoral dissertation. Amsterdam: University of Amsterdam.

McDonald, A.S. (2001). The prevalence and effects of test anxiety in school children. *Educational Psychology, 21* (2), 89-102.

Montessori, M. (1912). *The Montessori method.* New York: Frederick Stokes, Co.

Morra, S., Gobbo, C., Marini, Z., & Sheese, R. (2008). *Cognitive Development: Neo-Piagetian perspectives.* New York: Lawrence Erlbaum Associates.

Moyer, P. S. (2001). Are we having fun yet? How teachers use manipulatives to teach mathematics. *Educational Studies in Mathematics, 47*, 175-197.

Olkun, S. (2003). Comparing computer versus concrete manipulatives in learning 2D geometry. *Journal of Computers in Mathematics and Science Teaching, 22*, 43–56.

Olkun, S., Altun, A. & Smith, G. (2005). Computers and 2D geometric learning of Turkish fourth and fifth graders. *British Journal of Educational Technology, 36*, 317–326.

O'Malley, C. & Stanton-Fraser, D. (2004). Literature review in learning with tangible technologies. *Bristol: Nesta FutureLab Series,* report 12.

Papert, S. (1980). *Mindstorms: Children, Computers, and Powerful Ideas*. Basic Books, Inc.

Piaget, J. (1936/1977). *The origin of intelligence in the child*. London: Penguin books.

Piaget, J. (1952). *The origins of intelligence in children.* New York: W.W. Norton & Company, Inc.

Piaget, J. (1954/1971). *The construction of reality in the child*. New York: Ballantine Books.

Piaget, J.(1962). *Play, dreams and imitation in childhood.* New York: W.W. Norton & Company, Inc.

Piaget, J. (1976). *To understand is to invent.* The future of education. New York: Penguin Books.

Prensky, M. (2002). The motivation of game play: The real twenty-first century learning revolution. *On the Horizon*, *10 (1),* 5 - 11.

Price, S., Roger, Y. (2004). Let's get physical: The learning benefits of interacting in digitally augmented physical spaces. *Computers & Education, 43*, 137-151.

Ramsden, S., Richardson, F.M., Josse, G., Thomas, M.S.C., Ellis, C., Shakeshaft, C., Seghier, M.L. & Price, C.J. (2011). *Verbal and non-verbal intelligence changes in the teenage brain.* Nature Letters. Published online 19 October 2011.

Raven, J.C., Raven, J. & Court, J.H. (2004, Dutch translation 2006). *Manual for Raven's Progressive Matrices.* San Antonio, USA: Harcourt Assessment B.V.

Reid, S.E. (2001). Introduction: The psychology of play and games. In E.S. Charles & S.E. Reid (Eds.) *Game play: Therapeutic use of childhood games* (pp. 1–39). New York: Wiley.

Resing, W.C.M. (2006a). Using children's "ability to learn" in diagnosis and assessment. *Educational and Child Psychology, 23,* 6-11.

Resing, W.C.M. (2006b). *Zicht op potentieel. Over dynamisch testen, variabiliteit in oplossingsgedrag en leerpotentieel van kinderen.* Oratie. Universiteit Leiden: Maart 2006.

Resing, W.C.M., & Elliott, J.G. (2008). *Dynamic testing utilizing graduated prompts techniques and electronic scaffolding.* Paper presented at the APERA 2008 Conference on educational research. Singapore, 26-28 November 2008.

Resing, W.C.M., & Elliott, J.G. (2011). Dynamic testing with tangible electronics: Measuring children's change in strategy use with a series completion task. *British Journal of Educational Psychology, 81,* 579-605.

Resing, W.C.M., Steijn, W.M.P., Xenidou-Dervou, I., Stevenson, C.E. & Elliott, J. (2011). Computerized dynamic testing: a study of the potential of an approach using sensor technology. *Journal of cognitive education and psychology, 10* (2), 178-194.

Resing, W. C. M., Xenidou-Dervou, I., Steijn, W. M.P., & Elliott, J.G. (2011). A "picture" of children's potential for learning: Looking into strategy changes and working memory by dynamic testing. *Learning and Individual Differences.* doi.org/10.1016/j.lindif.2011.11.002.

Resnick, M., Martin, F., Berg, R., Borovoy, R., Colella, V., Kramer, K. & Silverman, B. (1998). Digital manipulatives: new toys to think with. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '98). ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 281-287.

Revelle, G., Zuckerman, O., Druin, A., and Bolas, M. 2005. Tangible user interfaces for children. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (Portland, OR, USA, April 02 - 07, 2005). CHI '05. ACM, New York, NY, 2051-2052.

Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research & Development, 44* (2), 43-58.

Roig, I., Resing, W. C. M., Herk, R. van, Meijles, J., & Verhaegh, J. (2008). *Dynamic testing: The use of a tangible table*. Poster presented at the International School Psychologist Assocation Conference, Utrecht, 8-12 July, 2008.

Rosen, S. (2003). Auditory processing in dyslexia and specific language impairment: is there a deficit? What is its nature? Does it explain anything? *Journal of Phonetics, 31* (3-4), 509 - 527.

Rugel, R.P. (1974). WISC subtest scores of disabled readers: A review with respect to Bannatyne's recategorization. *Journal of Learning disabilities, 7* (1), 48-55.

Salen, K. and Zimmerman, E. (2004). *Rules of play: Game design fundamentals.* Cambridge, Masssachusetts: The MIT Press.

Satyanarayanan, M. (2001). Pervasive computing: Vision and challenges. *IEEE Personal Communications 8* (4), 10–17.

Shaer, O. & Hornecker, E. (2010). Tangible User Interfaces: Past, Present and Future Directions. *Foundations and Trends® in Human-Computer Interaction*, *3* (1-2), 1-137.

Sherin, B., Reiser, B.J. & Edelson, D. (2004). Scaffolding Analysis: Extending the Scaffolding Metaphor to Learning Artifacts. *Journal of the Learning Sciences, 13* (3), 387–421.

Siegler, R.S., & Alibali, M. (2005). *Children's thinking.* Fourth edition. New Jersey: Pearson/ Prentice Hall.

Siegler, R.S., DeLoache, J., & Eisenberg, N. (2011). *How children develop*. New York: Worth Palgrave MacMillan.

Slater, A., Hocking, I., & Loose, J. (2003). *Theories and issues in development.* In A. Slater & G. Bremner, (Eds.), An introduction to developmental psychology. Oxford: Blackwell.

Smilde, A., Bro, R. & Geladi, P. (2004). *Multi-way analysis.* John Wiley & Sons.

Soute, I., Markopoulos, P., & Magielse, R. (2009). Head Up Games: combining the best of both worlds by merging traditional and digital play. *Personal and Ubiquitous Computing, 14 (5),* 435-444.

Sowell, E. J. (1989). Effects of manipulative materials in mathematics instruction. *Journal for Research in Mathematics Education, 20,* 498-505.

Sternberg, R. J., & Berg, C. A. (Eds.) (1992). *Intellectual development*. New York: Cambridge University Press.

Sternberg, R. J. & Grigorenko, E. L. (2002). *Dynamic Testing: The Nature and Measurement of Learning Potential.* New York: Cambridge University Press.

Stienstra, M. & Hoonhout, J. (2002) TOONS Toys. Interaction toys as a means to create a fun experience. In *Proc. of the Interaction Design and Children conference (IDC) 2002*, Eindhoven, Shaker Publishing BV, Maastricht, 199–210.

Swanson, H. L. (2006). Cognitive processes that underlie mathematical precociousness in young children. *Journal of Experimental Child Psychology, 93* (3), 239-264.

Sweller, J., van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Ter Braak, R. (2011). *Het stimuleren van de ontwikkeling van het begrip van plaatsbepalende voorzetsels en de ruimtelijke oriëntatie bij jonge kinderen.* Radboud Universiteit Nijmegen: Master thesis.

Terrenghi, L., Kranz, M., Holleis, P., & Schmidt, A. (2006). A cube to learn: a tangible user interface for the design of a learning appliance. *Personal and Ubiquitous Computing, 10*, 153-158.

Thurman, S.K., & McGrath, M.C. (2008). Environmentally based assessment practices: viable alternatives to standardized assessment for assessing emergent literacy skills in young children. *Reading & Writing Quarterly, 24,* 7-24.

Tobias, R. D. (1995). An introduction to partial least squares regression. In *Proc. of SUGI*, 1-8.

Tzuriel, D., & Shamir, A. (2002). The effects of mediation in computer assisted dynamic assessment. *Journal of Computer Assisted Learning, 18*, 21-32.

Tzuriel, D. (2001). *Dynamic assessment of young children.* New York: Kluwer Academic/Plenum Publishers.

Ulicsak, M. & Wright, M. (2010). *Games in Education: Serious Games*. Bristol: Futurelab. Retrieved October 4, 2012 from http://media.futurelab.org.uk/resources/documents/lit_reviews/Serious-Games_Review.pdf.

Ullmer, B. & Ishii, H. (2000). Emerging frameworks for tangible user interfaces. *IBM Systems Journal 39,* 915–931.

Valsiner. J. (2005). Scaffolding within the structure of dialogical self: Hierarchical dynamics of semiotic mediation. *New Ideas in Psychology, 23* (3), 197–206.

Van Herk, R., Verhaegh, J. & Fontijn, W. F. (2009). ESPranto SDK: an adaptive programming environment for tangible applications. In *Proc. of the 27th international Conference on Human Factors in Computing Systems* (Boston, MA, USA, April 04 - 09, 2009). CHI '09. ACM, New York, NY, 849-858.

Van Parreren, C.F. & Carpay, J.A.M. (1980). *Sovjetpsychologen over onderwijs en cognitieve ontwikkeling [Sovjet psychologists on education and cognitive development.].* Groningen: Wolters-Noordhoff.

Verhaegh,J., Fontijn, W., Aarts, E., Boer, L., & Van de Wouw, D. (2011). A development support bubble for children. *J. Ambient Intell. Smart Environ. 3* (1), 27-35.

Verhaegh, J., Fontijn, W.F.J., Aarts, E.H.L. & Resing, W.C.M. (2012). In-game assessment and training of nonverbal cognitive skills using TagTiles. *Journal of Personal and Ubiquitous Computing,* DOI: 10.1007/s00779-012-0527-0.

Verhaegh, J., Fontijn, W., and Hoonhout, J. (2007) TagTiles: optimal challenge in educational electronics. In *Proc. of the 1st international Conference on Tangible and Embedded interaction. TEI '07.* ACM, New York, NY, 187-190.

Verhaegh, J., Fontijn, W.F.J. & Resing, W.C.M. (submitted). On the correlation between children's performances on electronic board tasks and nonverbal intelligence test measures.

Verhaegh, J., Resing, W.C.M., Jacobs, A.P.A. & Fontijn, W.F.J. (2009). Playing with blocks or with the computer? Solving complex visual-spatial reasoning tasks: Comparing children's performance on tangible and virtual puzzles. *Educational & Child Psychology, Vol. 26* (3), 18-29.

Verhaegh, J., Soute, I.A.C., Kessels, A.C.J.M. & Markopoulus, P. (2006). On the design of Camelot, an outdoor game for children. In *Proc. of conference on Interaction Design and Children (IDC), Tampere, Finland, June 07 - 09, 2006*, (*pp. 9-16*). New York: ACM.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*, 250-270.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Wechsler, D. (1949). The Wechsler Intelligence Scale for Children. New York: Psychological Corp.

Wechsler, D. (2005). Wechsler Intelligence Scale for Children (WISC-III^{NL}), Handleiding en Verantwoording, Derde Editie NL. London: Harcourt Assessment.

Weiser, M. (1991). The computer for the twenty-first century. *Scientific American 165 (3)*, 94–104.

Wood, D. J., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychiatry and Psychology, 17* (2), 89-100.

Wijshoff, M. (2010). *Psychological assessment during games: review of literature and psychological tests.* Maastricht University: Internal student report.

Summary

Assessment and development of cognitive skills using tangible electronic board games

Serious games on the TUI TagTiles

When designing the educational tools and methods of the future, putting the child and its natural way of developing at the center offers great benefits. The child will be more motivated and at the same time the educational yield will be higher and more targeted. In this dissertation it is shown that electronic tangible systems like the TagTiles console can offer integral, personalized development of children in the areas of cognitive, fine motor and social skills for assessment, education and therapy, in a manner that builds on natural forms of play of children.

Tangible User Interfaces (TUIs) are potentially highly effective tools for education combining physical interfaces with computing power, enabling easy-to-use and robust applications that are enjoyable and motivating. The topic of this dissertation is whether and how TUIs can be developed that are effective for developing cognitive skills of children. Classical theories on cognitive development were used as a theoretical foundation for the development of a TUI-based educational application, such as the role of sensorimotor abilities for cognitive development as described by Piaget. Vygotsky's concept of the Zone of Proximal Development was used to inspire the implementation of adaptivity in the educational application. The research described consisted of three phases, each including an empirical study conducted at primary schools.

In the first phase the influence of the type of interface on the performance of children on an educational task was investigated. The use of a virtual, PC-based interface was compared to the use of a tangible, non-electronic interface for the same puzzle task. It was found that children (N=26, aged 5-7 years) were able to solve the tangible puzzle tasks on average almost twice as fast as the PC-based task, and needed considerably less instruction for the tangible version. The results of the study support the hypothesis that tangible interfaces offer a more suitable interface than a PC-based interface to educational tasks, at least for young children.

In the second phase it was validated whether a range of TUI-based tasks can be used to address nonverbal, cognitive skills. The applied tasks had been developed for use with 'TagTiles'. TagTiles is a tabletop electronic console with tangible game pieces developed by Serious Toys B.V. (www.serioustoys.com). The console includes a sensing board with an array of LED lights underneath and audio output. The system is controlled by manipulating game pieces on the TagTiles surface. Eight

visual-spatial tasks were created, intended to address different nonverbal cognitive skills such as (working) memory and spatial reasoning. Each task included abstract patterns consisting of colored tiles. For each task a different assignment is given to the player, such as mirroring the pattern, or repeating a sequence of tiles that lit up on the board. To validate which skills can be addressed with these tasks, children's performances on the TagTiles tasks were correlated with performances on several conventional psychometric instruments. This study included children aged 8-10 years and consisted of a pilot study (N=10) and an experiment (N=32). Significant correlations were found between the performances of children on the TagTiles tasks and the performances on nonverbal subtests of the Wechsler Intelligence Scale for Children III^{NL} (WISC-III^{NL}). Some tasks also showed significant correlations with Raven's Progressive Matrices, which is an intelligence test measuring deductive reasoning skills. The results of this study indicate that the developed tasks can be used to train skills that are measured in IQ tests.

In the third phase it was investigated whether the developed visual-spatial tasks kept their ability to address cognitive skills when embedded in a game. It was also tested whether children experienced this game, called 'Tap the little hedgehog', to be fun and intrinsically motivating. A fantasy theme was added to include the tasks in a natural way, to minimize and simplify the instructions needed to understand the game play and to make the tasks more fun to play. The difficulty of the task levels was made adaptive to the player's achievements. A reward structure was added to increase children's motivation to reach certain goals in the game as well as a support structure, created to help the child when needed, enabling independent play. The results of the empirical study (N=52, aged 7-9 years) with this game indicated that the added game context had not changed the essence of the tasks, as the performances were similar to those in the study in phase two. These findings support the hypothesis that TagTiles tasks applied in a game context can be used to assess and train a range of nonverbal skills.

For assessment purposes we concluded that TagTiles can be used to test at least part of the cognitive skills that are addressed with the applied conventional psychological measures, given the significant correlations that were found. Studies by others have shown that training of relevant skills such as working memory can improve aspects of intellectual functioning, in particular executive functioning and efficient use of working memory. This opens the exciting prospect that by practicing with TagTiles the performance on the mentioned skills may be enhanced, or that these skills may be more effectively used. This means that it would be useful to investigate whether, after further refinement and validation, the TagTiles tasks can be used for assessment and training of specific cognitive skills.

Based on the results of the conducted studies, it was concluded that the integral and personalized development of children in the areas of cognitive, fine motor and social skills for assessment, education and therapy can be facilitated with TUIs like TagTiles. Educational TUIs can profoundly change current education and assessment practices, offering an alternative that is enjoyable to the child and effective and accurate to the educational or assessment expert. The described way of creating a challenge using the Zone of Proximal Development can also be used to improve the experience with educational games.

Dankwoord

Vaak heb ik uitgekeken naar dit moment: het schrijven van het dankwoord. Omdat het proefschrift dan zo goed als af zou zijn, maar ook omdat het de kans geeft alle mensen te bedanken die eraan hebben bijgedragen. Het was niet altijd gemakkelijk, en het was erg fijn dat er op verschillende momenten mensen waren die mij hebben geholpen, geïnspireerd en gemotiveerd om verder te komen. Al deze mensen dank ik van harte. Een aantal wil ik graag apart noemen.

Ten eerste dank ik mijn co-promotor dr.ir. Willem Fontijn. Hij heeft mij in 2006 de kans geboden om mijn USI-afstudeerproject voort te zetten in de vorm van een PhD onderzoek bij Philips Research. Willem, ik ben je zeer dankbaar voor de kans die je mij gegeven hebt om dit project te mogen uitvoeren en nog veel meer voor je niet aflatende inzet gedurende de afgelopen jaren. Zelfs na je vertrek bij Philips en naast het opzetten van je eigen bedrijf Serious Toys ben je mij blijven begeleiden. Je vaak kritische maar altijd constructieve adviezen hebben mijn onderzoek naar een hoger niveau getild, terwijl je me tevens de ruimte gaf om me te ontwikkelen tot een zelfstandig onderzoeker. Daarnaast heb je ervoor gezorgd dat de uitkomsten van mijn onderzoek gebruikt worden in spelapplicaties voor de TikTegel. Daarmee is een belangrijke wens van mij uitgekomen.

Ten tweede dank ik mijn promotores prof.dr. Emile Aarts en prof.dr. Wilma Resing. Emile, hartelijk dank voor je begeleiding. Je hebt me zeer geholpen in het overzien en structureren van het project. Daarnaast ben ik je dankbaar voor je enthousiasme over mijn onderzoek en het vertrouwen dat je uitstraalde door regelmatig aan te geven 'dat het goed zou komen'. De gesprekken met jou heb ik altijd als zeer inspirerend ervaren.

Wilma, ook jou ben ik zeer dankbaar voor je bijdrage. Ik heb veel van je mogen leren over het testen en trainen van cognitieve vaardigheden bij kinderen, maar ook over het opzetten, uitvoeren en publiceren van wetenschappelijke studies. Ook heb je me geholpen in het vinden van studenten die konden helpen in de uitvoering van onderzoek, evenals psychologische testen of informatie daarover. Je hebt mijn werk altijd kritisch bekeken vanuit het perspectief van de psychologie, wat de kwaliteit van het proefschrift zeer ten goede is gekomen.

Ten derde dank ik dr.ir. Maddy Janse. Zij heeft mij vooral in de laatste maanden erg geholpen bij het schrijven, herzien en bediscussiëren van het proefschrift. Maddy hartelijk dank, ook voor je enthousiasme over mijn onderzoek en voor alle boeken die ik van je mocht lenen. Je hebt mij op verschillende momenten nieuwe energie gegeven om door te gaan.

Verder dank ik de leden van mijn promotiecommissie, prof.dr.ir. Caroline Hummels, prof.dr. Maartje Raijmakers, prof.dr. Huib de Ridder, prof.dr. Panos Markopoulos en prof.dr.ir. Aarnout Brombacher voor hun bereidheid zitting te nemen in deze commissie. Caroline Hummels dank ik ook voor de vergaderingen waarin zij mij samen met Kees Overbeeke advies heeft gegeven in de eerste jaren van mijn onderzoek.

Dr. Evelijne Hart de Ruijter-Bekker dank ik voor de periode waarin zij mij gecoacht heeft. Dankje voor je serieuze begeleiding en je advies om een plan te maken en daaraan vast te houden.

Ik dank de leden van de DQI groep bij Industrial Design van de TU/e voor hun gastvrijheid en behulpzaamheid, evenals de mensen van het Dynamic Testing Lab bij Onderwijs- en Ontwikkelingspsychologie van de Universiteit Leiden waarmee ik af en toe contact had.

Zonder medewerking van alle kinderen die hebben meegedaan aan mijn TikTegel studies op basisscholen en de daaraan voorafgaande pilot studies, was mijn onderzoek niet mogelijk geweest. Mijn dank is groot. Ook aan de ouders en docenten die hun medewerking hebben verleend.

Ik dank de managers van de afdelingen bij Philips Research voor de mogelijkheid die zij hebben gegeven om aan dit project te werken: Eelco Dijkstra, Jos van Haaren en Marieke van der Hoeven. Ook ben ik mijn Philips-collega's dankbaar die mij regelmatig in andere projecten 'ontzien' hebben op momenten dat er een PhD-deadline aankwam, evenals alle collega's die interesse in mijn onderzoek toonden.

In mijn PhD-werk heb ik hulp gehad van studenten die hun (afstudeer)projecten bij mij hebben uitgevoerd: Aljosja Jacobs, Laurens Boer, Julia Henning, Else Vink, Marèse Wijshoff en Joyce Krijgskamp, dank voor jullie hulp.

Doortje van de Wouw dank ik voor de ideeën die zij heeft gegeven bij het voorbereiden van het ontwerp van de omslag van het proefschrift.

Met de (oud)leden van de PPC commissie (de Philips-club van PhD studenten) heb ik veel lol gehad met de activiteiten die we samen organiseerden. Daarnaast zijn ook de serieuze gesprekken over het uitvoeren van een PhD binnen Philips waardevol voor mij geweest. Dankjullie wel!

Ook met mijn oud-kamergenoten Robert van Herk en Pavan Dadlani heb ik bij Philips veel plezier gehad. Robert, dankje, ook voor (je hulp bij) het programmeren van de TagTiles speltaken. Het was leerzaam voor mij als psycholoog om te zien hoe informatici te werk gaan. Vooral je opinie over flowcharts is mij bijgebleven. Pavan, thank you too for all the fun we had. I especially enjoyed the ID creation sessions with the three of us, where our ideas often balanced between brilliance and nonsense.

Tot slot prijs ik mij zeer gelukkig met mijn lieve familie en vrienden. Ik ben jullie erg dankbaar voor de interesse die jullie de afgelopen jaren in mijn werk hebben getoond, maar vooral voor de steun die jullie mij bleven geven. Pap en mam, en Ben en Helga, ik dank jullie in het bijzonder, ook voor de menige uurtjes die jullie hebben opgepast zodat ik kon schrijven.

Anke en Rob, dank dat jullie mijn paranimfen willen zijn. Jullie interesse in mijn onderzoek en de gesprekken die we erover hadden, zijn belangrijk geweest voor mij.

Lieve Taeke, jij hebt heel wat offers gebracht om mij dit werk te kunnen laten afmaken. Ik ben je zeer dankbaar voor het geduld dat je hebt weten op te brengen en voor je niet aflatende steun. Jij hebt, in het bijzonder na de geboorte van Siem, ervoor gezorgd dat ik de motivatie en de rust heb hervonden om het te kunnen afmaken. Ik draag dit proefschrift dan ook aan jou op. Ik kijk ernaar uit om samen met jou en Siem weer heel veel leuke dingen te gaan doen.

Curriculum Vitae

Janneke Verhaegh was born on August 10, 1979 in Wageningen, The Netherlands. After finishing her secondary education (VWO) in Venlo in 1998, she moved to Maastricht to study at the Maastricht Academy of Fine Arts and Design. After one year she left the academy to study Psychology at the Maastricht University. She obtained her Master degree in Biological Psychology in 2004, after completing a graduation project at Philips Research in Eindhoven. The project focused on personality expression in robots and the perception of robotic personalities by children. This project led to Janneke's application for the User-System Interaction (USI) programme at the Eindhoven University of Technology. After graduating from this progamme in 2006 within a project that had resulted in the first prototype of TagTiles, Janneke started as a PhD student at Philips Research to continue her research on skill development through TagTiles games. This dissertation is the result. Since 2010 Janneke is employed by Philips Research as a Research Scientist.

List of Publications

Journal articles

Janneke Verhaegh, Wilma C.M. Resing, Aljosja P.A. Jacobs & Willem F.J. Fontijn (2009). Playing with blocks or with the computer? Solving complex visual-spatial reasoning tasks: Comparing children's performance on tangible and virtual puzzles. *Educational & Child Psychology, 26* (3), 18-29.

Verhaegh, J., Fontijn, W., Aarts, E., Boer, L., & van de Wouw, D. (2011). A development support bubble for children. *J. Ambient Intell. Smart Environ. 3* (1), 27-35.

Verhaegh, J., Fontijn, W. F. J., Aarts, E.H.L. & Resing, W.C.M. (2012). In-game assessment and training of nonverbal cognitive skills using TagTiles. *Personal and Ubiquitous Computing. Special Issue on Child Computer Interaction;* Epub ahead of print. Springer. DOI 10.1007/s00779-012-0527-0.

Janneke Verhaegh, Willem F. J. Fontijn & Wilma C. M. Resing. On the correlation between children's performances on electronic board tasks and nonverbal intelligence test measures. (submitted).

Henning, J.R., Verhaegh, J. & Resing, W.C.M. (2011). Creating an individualized learning situation using scaffolding in a tangible electronic series completion task. *Educational and Child Psychology, 28* (2), 85-100.

Peer-reviewed conference contributions

Verhaegh, J., Fontijn, W., & Hoonhout, J. (2007). TagTiles: optimal challenge in educational electronics. In *Proc. of the 1st international Conference on Tangible and Embedded interaction* (Baton Rouge, Louisiana, February 15 - 17, 2007). TEI '07. ACM, New York, NY, 187-190.

Verhaegh, J., Hoonhout, J., & Fontijn, W. (2007). Effective use of fun with a tangible interaction console. In *Proc. of the 4th International Symposium on Pervasive Gaming Applications* (Salzburg, Austria, June 11.-12, 2007) PerGames 2007, Shaker-Verlag, Aachen, 177-178.

Verhaegh, J., Fontijn, W., & Jacobs, A. (2008). On the benefits of tangible interfaces for educational games. In *Proc. of Digitel 2008* Banff, Canada, Nov. 17-19, 2008, IEEE Computer Society and Conference Publishing Services (CPS), 141-145. (Awarded Best Short Paper)

Verhaegh, J., Fontijn, W., Aarts, E., Boer, L., & van de Wouw, D. (2009). Creating a Development Support Bubble for Children. In *Proc. of the European Conference on Ambient intelligence* (Salzburg, Austria, November 18 - 21, 2009). Lecture Notes In Computer Science, vol. 5859. Springer-Verlag, Berlin, Heidelberg, 186-195.

Van Herk, R., Verhaegh, J., & Fontijn, W. F. (2009). ESPranto SDK: an adaptive programming environment for tangible applications. In *Proc. of the 27th international Conference on Human Factors in Computing Systems* (Boston, MA, USA, April 04 - 09, 2009). CHI '09. ACM, New York, NY, 849-858.

Hendrix, K., van Herk, R., Verhaegh, J., & Markopoulos, P. (2009). Increasing children's social competence through games, an exploratory study. In *Proc. of the 8th international Conference on interaction Design and Children* (Como, Italy, June 03 - 05, 2009). IDC '09. ACM, New York, NY, 182-185.