The handle http://hdl.handle.net/1887/25194 holds various files of this Leiden University dissertation.

Author: Chabani, Elaheh (Ellahe)
Title: Enhancing visuospatial processing skills in children
Issue Date: 2014-04-15
Chapter 5

General Discussion
Visuospatial processing skills (VSPS) is the over-arching concept I used to refer to a set of skills in gathering visual information from the environment and integrate them to derive meaning from what one sees. VSPS is required for successful academic performance and for many activities in daily living. While there is evidence of the trainability of VSPS in adolescents and adults (e.g., Newcombe & Frick, 2010; Uttal, et al. 2012), not much is known about school age children. Moreover, despite research on spatial intelligence in school age children, spatial content is hardly considered in school activity and there is hardly any awareness of teachers that VSPS are important (e.g., Krakowski, Ratliff, Gomez& Levine, 2010). The main goal of the present project was to develop instruments and methods for assessing and training VSPS in school age children, and to stimulate interest in elaborating and advancing research on VSPS in researchers. Toward this end my colleagues and I have developed a computerized game-based instrument: the “TangSolver”. Studies presented in this thesis described the rational, development, and implementation of this instrument. In the present chapter, I will summarize my general findings and discuss the in relation to limitations of my research and its implications for education, clinical practice, and research.

**Summary of findings**

This project utilized a pretest-intervention/training-posttest design. The pre- and post-tests did not include feedback, which make them comparable to conventional VSPS tests. Between pre- and post-test, the experimental group received training and feedback by means of either face-to-face or computerized instruction. The performance change in the child’s ability in solving visuospatial tasks from pre- to post-test indicate the trainability of VSPS (how much can be learned from short intervention) and as well the efficacy of the training. The first step taken was the evaluation of our manipulative experimental material in typical children (see chapter 2). The second was the evaluation of its computerized version (see chapter 3), followed by comparisons of its effectiveness in typical developing children (TD) and children with ASD (ASD) (see chapter 4). With respect to trainability of VSPS the following conclusions can be drawn.

First, the greater gain observed in the experimental groups that received training than in the control group that did not receive training (see chapters 2 & 3) provides evidence of VSPS trainability in typically developing children and demonstrates the efficacy of the interventions.

Second, a particular interesting result was the trainability of ASD children reported in chapter 4. It has been frequently reported that individuals with ASD exhibit superior abilities to identify fine stimulus features in spatial tasks (Caron, et al., 2006; Mottron, et al., 2003; Shah & Frith, 1993), but are limited in their ability to derive organized wholes from perceptual parts. This atypical ability has been linked to their limited use of gestalt grouping heuristics and/or the failure to consider the entire visual context (Happe, 1996). More recently, Happe and Frith (2006) have suggested that the weakness of global processing can be overcome in tasks with explicit demands. Even though in our experiment ASD participants were not highly functioning and even though we did not stress explicit task demands in terms of local and/or global
processing, we could not find any evidence that ASD children might be systematically impaired with respect to either global or local processing. The few differences we observed were all eliminated by training, which supports suggestions that task understanding can be improved through training (Hill & Frith, 2003). Furthermore, many have stressed that learning in individuals with ASD is compromised by a restricted range of interests, inflexibility to a non-functional routine, resistance to change, and deficiencies in reciprocal social interaction (Richler, Bishop, Kleinke & Lord, 2007; Wolfberg, 2009). Yet, ASDs involved in our experiment did not show specific resistance to task accomplishment, suggesting that apparent resistance to learning in conventional ways might be more depending on interest and motivation rather than intrinsic capacity. As outlined by some authors (e.g., Roelfsema, 2006; Gilbert & Sigman, 2007), tasks that rely on grouping of elementary features require both levels of Gestalt rules, low-level like similarity, and high-level of grouping cues, such as familiarity with the shape. We consider it possible that the items used in our testing and training procedures played an important role, which calls for further experiments with more abstract items.

Third, to assess the potential of the current computerized VSPS training vis-à-vis traditional face-to-face training, we compared the effectiveness of these two trainings methods. We considered the computer training more suitable for those interested in autonomous learning, while face-to-face training was more personalized by employing manipulative material. Computerized training made use of visual hints (e.g. visually scaffolding, full or partial solution showing etc.), while face-to-face training employed the usual mix of verbal and visual cues. Irrespective of all these differences, the two training methods produced a coolant outcomes in both normally developing and ASD children. These observations provide evidence for the value of purely visual instruction cues (Mayer & Moreno, 2003; Mayer, 2009), at least in tasks like ours. This provides interesting avenues for the training of children with language-based learning disabilities (e.g., Guarnera, Commodari & Peluso, 2013; Dalton & Proctor, 2007; Newhall, 2012; Smith & Tyler, 2009).

Fourth, training effects in simple transformations tasks were equivalent in all three groups, and thus independent of actual practice. In contrast, there were specific training effects of the complex transformations task, which required flipping or rotating MPs. This suggests that simple transformations tests might be more appropriate for younger children or children with impaired mental rotation abilities (e.g., Guarnera, Commodari & Peluso, 2013), while the complex transformations task is more diagnostic with respect to VSPS proper.

Taken together, I am confident in concluding that even short training (Uttal et al., 2012) can enhance VSPS considerably in both TD and ASD children. It is reasonable to consider that more intense training might results in even higher gain. While computer training can be as efficient as face-to-face training (e.g., Koppenhaver & Erickson 2003; Pennington, 2010; Ramdoss, et al., 2011), the added value is to reduce barriers such as complexity of the interventions in term of materials and resources, and time required for implementing (Hale, et al, 2010). Moreover, training through computer applications does not
rely on teacher expertise. While the material developed for this study was suitable for our purposes, more items will be necessary to extend training to other age groups and populations. Moreover, more tests would be helpful to predict pre-test performance and training efficiency, to define the most influential and efficient types of feedback, to diagnose which strategies learners are using, how strategy choice can be guided by feedback, and follow-up studies will be necessary to evaluate the longer impact of VSPS training.

**Future direction**

The findings of the current research raise further questions regarding data analysis and design, which may be addressed in future research projects.

**Pre- and post-testing evaluation.** For practical purposes, I have analyzed my data with the most widely used measurement model in education, which is using the logic of True Score Theory (TST) to assess change from pre- to post-test by means of ANOVAs for repeated measures. Even though TST is mathematically simple and straightforward, it has some drawbacks however. For example, TST is oriented towards the total test score rather than the individual items, subject characteristics, and test features, based on the assumption that the total (observed) score on a test is the ‘true score’ plus some random error which is assumed to be the same across all individuals (e.g., de Klerk, 2008; Kline, 2005). Many authors have claimed that in TST a higher score does not necessarily mean that the person has more ability with respect to the assessed trait than a person scoring lower; that is, a person scoring 80% does not necessarily have more ability than a person scoring 60% (e.g., de Klerk, 2008; Kline, 2005; Shultz & Whitney, 2005). To address that issue, Item Response Theory (IRT) and the Rasch model (a special case of IRT model) have been suggested as an alternative or supplementary model to TST (e.g., Bechger, et al. 2003; Bond & Fox, 2007; Wu & Adams, 2007). IRT is based on the premise that a test taker’s performance on a given item is determined by two factors: the test taker’s level of ability and the characteristics of the item. Others have suggested a generalized mixed-effects regression model (West, et al. 2007) that is quite robust to missing data, but the disadvantage is that such models are computationally complex. Regardless of the particular statistical method, a critical issue is how to get a better insight into the process of improving on a task or ability (e.g., Romero, Ventura & Garcia, 2008), which calls for a more detailed analysis of the training data. Furthermore, informal observations during the study suggested that some weaker students who initially were relying on extra help, were managing well after few trials—an improvement that apparently was lost until the posttest was taken. Reversely, successful skill acquisition might need some time to consolidate, which may suggest that longer delays between pre- and post-test provides more reliable results. However, the testing schedule will often depend on pragmatic issues constrained by school activities, curricula, and the research timeframe, which may underestimate true learning effects. One interesting option is computer adaptive testing (Frick, 1992; Lee & Weiss, 2010), which however has the downside that the calibration of the item pool requires extensive data collection - pre-administered to a sizable sample prior to test development (Parshall, Spray, Kalohn & Davey, 2006). Such calibration requires high collaboration.
among researchers, teachers, and software designers, which is the main constraint but it is hoped to evolve in future. It should be kept in mind that quality control, item bank development and psychometric processes involved in construction task like the one used here is more complicated than designing multiple choice items.

**Differencing Methodology.** Dellow (2010) argued that experiments employing pretest, an instructional period, and an eventual posttest with identical or nearly identical items do not allow for the efficient detection of variance in learning. Dellwo suggests evaluating the initial and final levels during the training phase, which makes for a multi-stage pre/post design (see fig.1). For instance, the current training structure consists of starting with easiest level (4MPs) to hardest (7MPs) of the same figure, independently of the individual’s initial level. In multi-stage pre/post testing, the child’s initial level could be defined as the level that the child can manage the task easily at. Accordingly, the child could start at any level and move to an easier one if s/he cannot manage within one minute, until the “initial level” has been identified. From this initial level, the training would start up to the compilation of the last, most difficult level (7MPs in our case). The difference between initial level and finally achieve level would then serve to diagnose individual differences, which could be used to predict potential learning benefits.

**Figure 1.**
The first stage of a two-stage assessment scheme is bracketed by pre- and post-instruction tests T0 and T1. The second stage is bracketed by T1 and T2. The diagnostic tests are identical or nearly identical instruments designed to assess learning of key skills and concepts. From Dellow, (2010). Course assessment using multi-stage pre/post testing and the components of normalized change. Directedly reprinted with permission

**Independent variables.** The three variables included in our data analysis were time on task, Accuracy and Tasks completed. In contrast to Accuracy and Tasks completed, Time on task has produced disappointing results, thus we consider it unlikely that this variable is systematically related to those aspects of performance we were interested in. Possible solutions are either to increase the allocated time per items or to drop the time limits altogether. Furthermore, it is becoming more and more common for
computerized games to provide large varieties of measures i.e., number of in/correct pieces moved, time from when the first move was initiated after a figure presentation, and so forth. Such outcomes might bring more insight into performance variability, especially if combined with mathematical models of educational data mining (e.g., Aleven et al. 2010; Baker et al., 2010; Feng & Heffernan, 2010; Shute & Zapato-Rivera, 2012). Indeed, a more continuous process analysis of the training/intervention phase should be the eventual goal of assessment (see, www.ed.gov/technology and http://myweb.fsu.edu/vshute/publications.html). However, the analyses of training data were beyond the scope of this PhD project, not the least because of an urgent need to rapidly turn empirical findings into significant insights that guide teachers in their teaching strategies (Bouchet, Harley, Trevors & Azevedo, 2013). Nevertheless, process analyses will be an important next step (Shute & Zapato-Rivera, 2012; Gobert et al. 2012). To conclude, such spatial training might be a great tool for educators, but there is still much to be learned in terms of interpretation of data and the significance of these.