# Physical Touch-Based Rotation Processes of Primary School Students

Sven Bertel[1](✉), Stefanie Wetzel[1], and Steffi Zander[2]

[1] Flensburg University of Applied Sciences, Flensburg, Germany
{sven.bertel,stefanie.wetzel}@hs-flensburg.de
[2] Bauhaus-Universität Weimar, Weimar, Germany
steffi.zander@uni-weimar.de

**Abstract.** We present a novel method for deriving solution strategies used when solving rotation tasks. Process-based findings about physical rotation are obtained from an analysis of how angular disparity between stimuli changes over time. Data on angular disparity was gathered through a study on mental and physical rotation with 37 primary school students between the ages of 8 and 11. For controlling physical rotation, students used touch-based input on our iOS app *Rotate it!*, which also logged their interactions. Data on changes of angular disparity was used in the construction of Markov models. The models were employed to generate sets of synthetic angular disparity time courses, based on which we identified three distinct rotation-based solution strategies. Our analysis has implications for understanding processes involved in physical and mental rotation alike. It helps to lay grounds on which novel interactive diagnostic and training tools for spatial skills can be developed.

**Keywords:** Physical and mental rotation · Solution strategies for spatial tasks · Spatial skills · Qualitative description of rotation processes

## 1 Introduction

Spatial intelligence is a main part within multi-component models of intelligence [11], important for performance in many visual and spatial tasks. It is of special importance for tasks within STEM domains [43, 46]. Not everyone is equally good at spatial and visual tasks, though. Among other factors, individual performance has been strongly linked to training and practice (e.g., [3, 40]), which may affect the availability and use of mental strategies, or simply lead to more effective mental processing (e.g. [20]). In spite of this, the training of spatial skills is widely underrepresented in educational curricula at primary and secondary levels when compared to how verbal or mathematical skills are trained (e.g., for schools in the U.S., [5]). It has been pointed out that enriching school curricula to include an adequate training of spatial skills will likely increase students' professional participation in STEM domains [44]. More generally, investing into cognitive

skill formation has been linked to economic success in life for the individual, as well as for the society as a whole [18].

This situation opens up opportunities for educators and spatial cognition researchers alike. Aspects that need to be addressed both theoretically and practically relate to which spatial tasks should best be trained when, with whom, in which form, and at which age (cf. [35]). The crucial aspect is what precisely constitutes an "adequate training of spatial skills". A number of previous studies have focused on general issues of training, such as on how narrowly or broadly expertise in one spatial skill transfers to other spatial skills [44,50]. Population subgroups for which specific spatial ability levels have been investigated have been typically chosen based on age (mostly focusing on developmental aspects, [9,33,35]) or sex/gender [22,31]. Less research exists that puts an emphasis on individual spatial abilities and on differences in the mental strategies that are employed by different individuals (see e.g. [13,28], for recent exceptions).

As spatial abilities, skills, and strategies frequently differ inter-individually, it seems plausible to assume that the best way to practically train individual spatial abilities in a student population through spatial and visual tasks will also involve different training methods for different individuals. The general aim of this contribution is to lay suitable empirical and model-based grounds on which effective individualised training programs may subsequently be developed.

## 1.1   Foci of this Contribution

In this contribution, we will focus on procedural aspects of rotation tasks, which were registered through touch events on mobile devices, and will use physical rotation as a proxy for mental rotation. Our questions include: How do typical solution processes differ between correct and incorrect trials? How do they differ for different tasks? What differences can be observed between different, but similarly successful solution processes for the same task? Novel, model-based descriptions of solution processes will be employed to derive information about successful and unsuccessful strategies.

## 1.2   Outline

In the following section, we will establish a focus on mental rotation as one class of generally well-researched spatial tasks that, in spite of much research, still offers a number of interesting research opportunities. We developed the iOS app *Rotate it!* to present spatial tasks to users and to log their interactions while solving rotation tasks. In Sect. 3, we will describe an exploratory study with students at a primary school in which we employed *Rotate it!* to record solution processes of students' physical rotations of objects. Importantly, we will make the case that physical rotation trajectories may be used to gain novel, process-based insights also into students' general and individual mental rotation strategies. Next, we will analyse and discuss results, in large parts based on qualitative, probabilistic process models built based on the data that we gathered during the study. As a result of this modelling step, a population of synthetic trajectories

will be generated that reflect frequent problem solving moves. Based on these, three distinct rotation strategies will be identified and compared on success rates and times per task. We will conclude by drawing implications for further process-based research into spatial problem solving as well as for a better description of individual spatial abilities, and will discuss ramifications for developing more effective and efficient individualised, dynamic training programs for spatial skills.

## 2    Mental and Physical Rotation

Tasks that have been frequently employed to study various factors of spatial intelligence include, among many others, mental transformations of perceived or imagined objects (e.g., mental rotation, [38,45]) and mental transformations of perceived or imagined scenes (e.g., mental orientation or perspective taking, [30]). While underlying mental skills for both groups of tasks seem related and performance is usually found to be significantly correlated, there is nevertheless evidence to assume a fundamental dissociation of faculties [30].

The two most frequently used mental rotation test paradigms employ 2D visualisations of 3D objects made by joining cubes side by side (cf. Fig. 1). An individual task either consists of two object visualisations (following [38]) or five (following [45]). In the former case, participants need to decide whether the two visualisations do, or do not, show the same object. In the latter, they need to compare the first visualisation to the four other ones and decide which two of the four show the same object as the first.

We define physical rotation tasks as tasks during which the problem solver has an option to physically rotate object representations before deciding whether same or different objects are shown. This definition of physical rotation is similar to how the term was recently employed by Gardony et al. [12].

### 2.1    Similarities, Commonalities, and Analogies

There is robust evidence to assume that, in many cases and for most participants, mental rotation solution processes will involve some form of visual mental imagery (for exceptions, see e.g. [32]). The processes are thought to be similar to those active during the physical rotation of objects. Support for this assumption comes, first, from behavioural studies: For pairs of visualisations that show the same object, correct response times are linearly proportional to the angular offset between the two object visualisations (angular disparity effect; [6,38]). The existence of the angular disparity effect has been often interpreted as a sign that, procedurally, mental rotation is analogous to physically rotating an object. In this sense, mental rotation can be seen as a kind of mental simulation that is comparable to a physical rotation process and during which the two visualisations become gradually aligned.

Further support for such interpretation comes from dual-task studies, during which participants generate additional physical, rotary movements with their arms or hands. When the direction of the physical rotation corresponds to the

direction of the mental rotation, mental rotation response times are decreased; when the directions do not correspond, response times are increased [47,49]. Enacting (congruent) physical gestures has been found to improve performance in spatial visualisation tasks, such as during mental rotation, especially in people who have difficulties at such tasks [4]. Explanations for the beneficial effect of gesturing do not just point at an offloading of mental spatial representations, but to a general improvement of mental spatial transformations through using gestures [4]. In a recent comparison between mental and physical rotation tasks (the latter being controlled through rotating a ball in hand), Gardony et al. [12] have found comparable angular disparity effects for both mental and manual rotation.

Secondly, support for common (or, at least, overlapping) mental processes in mental and physical rotation comes from developmental perspectives (cf. [35]). Mental rotation task performance in children at the age of 5 to 6 years correlates with individual motor control (for rotation) [23], to the extent that children with impaired motor control have also shown reduced mental rotation performance (e.g., for overweight children [24]).

Thirdly, there is a body of neuroanatomical evidence for common or overlapping functions in mental and physical rotation, or for a mental simulation of a physical process during mental rotation. For example, areas within the primary motor cortex have been found to be active also during mental rotation [7,10]. Finding M1 activation was dependent on whether participants had been instructed to imagine that it was them who were rotating an object (activation) or that the object was being rotated by some exogenous force (no activation; [29]).

## 2.2   The Training of Mental and Physical Rotation

We have discussed in the introduction that, for spatial and visuo-spatial tasks in general, performance is strongly related to practice and training. More specifically, this statement also holds for mental rotation tasks (e.g., [20,48]), for which effects of training can last up to several months [40]. Importantly, mental rotation skills can also be trained by manual rotation, while the converse is not true: manual rotation skills are not significantly improved by mental rotation training alone [1]. Effects of practice seem to be process-based, rather than instance-based, and there is evidence that mental rotation skills can transfer to other spatial tasks, such as to paper folding [50].

Various studies have investigated mental rotation abilities in infants, preschoolers, and students at primary and secondary schools [9,21]. Mental rotation of some sorts can be reliably performed around the age of 4 to 5 years [33], while mentally rotating 3D objects is still difficult at an age of 8 to 10 [25]. Neurologically, structures involved in spatial thinking are fully functional already at a very early age, so that, again neurologically speaking, early practice seems important, if not essential, for developing high spatial thinking skills later on [15].

Spatial task performance of both males and females benefits from practice and training. Training may in fact reduce, or even fully level, existing gender

differences [8,42]. While older research (including meta-analyses) consistently reported gender differences for paper-and-pencil-based mental rotation tasks to the advantage of males [31], more recent studies have failed to detect a gender difference for all types of mental rotation tasks. Notably, no difference was found for stimuli that employ 3D cube figures, while stable differences were found for stimuli based on 2D polygonal objects [22]. Using human figures as test stimuli decreased gender differences in mental rotation scores compared to using 3D cube figures [2]. For test participants around the age of 7 to 12, the extent of an existing male benefit in mental rotation varies strongly with the selected stimulus type [37].

## 2.3   Strategies, Theories, Models, and Research Methods

Early, defining procedure-oriented insights into mental rotation were provided by Shepard and Metzler [38], and Cooper and Shepard [6] through describing the angular disparity effect: Response times increase in a linear fashion with the angular disparity between the object visualisations. In an eye tracking study (and based on previous, conceptual suggestions by Metzler and Shepard [34]), Just and Carpenter [26] identified three major, idealised phases of mental rotation: (1) *search*, during which pairs of superficially corresponding object segments are identified; (2) a series of *transform-and-compare* operations. During each operation, two corresponding segments are first rotated in stepwise fashion such that they gradually become more aligned. After each transformation, a check for sufficient congruence is performed; and (3) *confirmation*, during which an additional check is performed whether the transformation of the matching segments has also brought the objects as a whole into sufficient congruence. According to Just and Carpenter [26], individual transformations in phase 2 are in 50° increments and a 25°-offset constitutes a suitable level of congruence between two objects to stop repeating the phase.

Khooshabeh et al. [28] examined individual differences in mental rotation strategies. In particular, they focused on whether participants employed piecemeal rotation strategies similar to the three-phase strategy described above, or holistic strategies in which the entire object is mentally rotated. Strategy choice was found to depend on individual working memory capacity, such that lower capacity was associated with using piecemeal strategies. In terms of adults' strategy use for rotating 2D polygons, holistic strategies have more frequently been found with males and piecemeal strategies with females [19]. It is important to keep in mind that rotation-based strategies, whether piecemeal or holistic, are not the only type of strategy for effectively solving mental rotation tasks. Other existing strategies may be more analytical and, for example, rely on comparing cube counts between roughly corresponding object segments. Logie et al. [32] reports on individuals who employed such analytic strategies in an fMRI study. It should be clear, that, when analytical strategies are used for solving mental rotation tasks, the mental processes involved are likely to be very different from mental processes involved in solving physical rotation tasks. Furthermore, people

who mentally rotate stimuli (*rotators*) have been associated with using holistic strategies, while *nonrotators* use analytic strategies [14].

Recent studies replicated the angular disparity effect as an increase of response time with angular disparity, though the found relationships were not always perfectly linear (see e.g., [12,17]). Naturally, an angular disparity effect as defined by Shepard and Metzler [38], and Cooper and Shepard [6] and used by Just and Carpenter [26] only exists for *same* stimuli, that is, when both object visualisations show the same object (cf. results of Gardony et al. [12], and Jansen and Heil [22]). For *different* stimuli, there likely exist multiple ways of trying to match visualisations without succeeding, while there usually exists only one perfect match between *same* stimuli.

It is interesting to note that very few approaches in the mental rotation literature have attempted to map involved mental processes to the degree attempted by Just and Carpenter [26]. While more recent contributions such as Khooshabeh et al. [28] shed light onto use of overall strategies, although on an individual problem-solver's basis, there is little detailed information provided as how to the solution processes progress over time. For example, how does the angular disparity change during the course of solving a mental rotation problem? How well does the model suggested by Just and Carpenter [26] really match stepwise mental rotations? The lack of research on this point is not surprising, as tracking a (hypothesised) mental process (e.g., of a changing angular disparity) over time is not easy. Likely, no current method based on fMRI or EEG will permit the tracking of transformations of mental representations with adequate temporal or spatial precision, nor does a suitable experimental methodology seem to exist for such an endeavour. However, the various similarities between mental and physical rotations discussed above may provide us with a window onto some of the inner workings of mental rotation.

In recent research, Gardony et al. [12] first opened up this window by contrasting mental and physical rotation tasks (inputs to the latter were through turning a ball held in hand). Stimulus objects were 3D cube figures taken from Peters and Battista [36]. Based on the tracking of angular offsets between stimuli over the course of solving a physical rotation task, Gardony et al. [12] established that, for successfully solved *same* tasks, participants on average rotated until a characteristic offset of around 30° was reached. This mark is surprisingly similar to the 25°-mark postulated by Just and Carpenter [26] as the final offset for transformations in mental rotation. In terms of method, the present contribution draws inspiration from Gardony et al. [12], though our analysis of time courses of angular offsets goes significantly beyond that. In addition, differences exist in the method of rotational input and in the targeted group of participants.

## 3   Our Study

Tablets are widely available in educational contexts today and allow for touch-based inputs. We decided to use a tablet-based app to present visuo-spatial training tasks to school students and to track their problem solving actions.
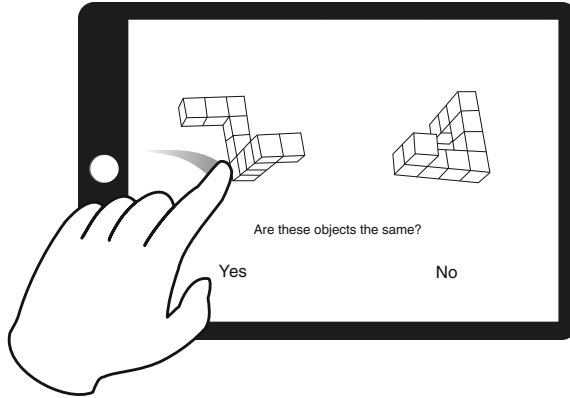
**Fig. 1.** A sample rotation task as presented in the *Rotate it!* app.

To this end, we developed the iPad app *Rotate it!* and employed it to track students' 2D touch gestures used for rotational control during the physical rotation of stimuli. *Rotate it!* displays a sequence of rotation tasks with two object visualisations each and permits rotating the left-hand figure using an Arcball interaction metaphor ([39]; see Fig. 1). We chose the Arcball, as it is easy to understand and does not require much familiarisation. All touch-based interactions (e.g., finger positions or object rotations over time) as well as response times and given answers are automatically logged.

The school curriculum at primary schools in the German state of Thuringia includes at least simple spatial tasks, such as the mental folding of cube nets [27,41]. We teamed up with a local primary school to conduct a pilot study that was aimed, first, at testing the *Rotate it!* app and, secondly, at gathering data about how students at the school solved rotation tasks. Mental and physical rotation tasks were administered in a within-subject design. Written parental and school approval for the study and for gathering data was sought and obtained.

During the study, students were presented with series of cube figure pairs (in the style used by Shepard and Metzler [38]) and had to decide for each pair whether the shown figures were identical, or not (see Fig. 1, for an illustration). In order to better compare our results with those of past research, we chose to employ the pairwise figure comparisons from the Vandenberg and Kuse Mental Rotation Test [45]. Reconstructed figures were used and, based on the 20 1-to-4 comparisons of the original instrument, we constructed 80 1-to-1 comparisons, which were grouped into two sets of 40 comparisons each. Each set contained 20 tasks with identical stimulus pairs (*same*, with initial angular disparities between 0° and 180°) and 20 tasks with *different* stimulus pairs. For the *physical* condition, tasks were presented via *Rotate it!* to allow us to log all actions that students produced. For the *mental* condition, we used a static paper-based version of the tasks, again with the intention of staying close to formats of original paper-based mental rotation test instruments. Stimulus size was kept constant

between conditions, and sheets in the *mental* condition had the same size and aspect ratio as the iPad screen in the *physical* condition. The order in which the two conditions were presented was balanced across all students to compensate for potential learning effects. Consequently, students either started with the iPad-based *physical* condition, followed by the paper-based *mental* condition, or vice versa. For the *mental* condition, time was only measured over all tasks due to practical constraints within the classrooms, and not for single tasks.

## 4    Results

We gathered complete data from 37 students (m = 19 / f = 18; age: 8–11, mean: 9.08 years). In a first step of analysis, we used a mixed within- and between-subjects design to compare data on mental effort, motivation and performance data between conditions. We found descriptive (though non-significant) differences in means for success (*mental*: 66.04%; *physical*: 71.13%), time per task (*mental*: 12.61 s; *physical*: 9.13 s), and mental efficiency (computed as a ratio of mental effort over time, with self-assessed ratings of mental effort; *mental*: 10.94; *physical*: 14.54). The overall success rate of 68.59% indicates that our tasks were suitable for our sample of students, as they were of medium to high task difficulty. For a detailed analysis and discussion of success rates, response times, and different efficiency measures, as well as of condition order, we refer to Zander et al. [51]. For the present context, we will focus on an analysis of changes of angular offsets gathered over the courses of physical rotation tasks.

**Courses of Angular Offsets.** For an initial analysis of rotational offsets, we pooled all the data that we had gathered through *Rotate it!*. We first wanted to see if our study qualitatively replicated the final angular disparity patterns for physical rotation found by Gardony et al. [12, p. 610]. This was of particular interest, as our input of rotational control differed from that in Gardony et al. [12] (touch and Arcball vs. rotating a ball in hand). As Fig. 2 shows, this was indeed the case: For *same* correct answers, we found an overall decrease towards low final angular disparities, which was larger for larger initial angular disparities. The final median angular disparity was 29° (N = 416). For *same* incorrect trials, the median angular disparity was 121° (N = 118). The difference to correct trials was significant (Z = −11.29, p < 0.001, r = 0.49). We focused our present analysis on *same* trials, as, in a first step, we are chiefly interested in characteristics and variation of successful problem solving strategies. Also, one has to note that the *different* trials derived from the Vandenberg and Kuse instrument [45] do not differ systematically (e.g., are mirror images, as in Gardony et al. [12]), but in many different ways. Consequently, data from *different* Vandenberg and Kuse pairwise comparisons cannot be easily aggregated across tasks.

**Angular Disparity Effect.** We were further interested in whether our data showed a linear angular disparity effect, similar to the one reported in the literature [6,38]. For *physical* rotation, initial angular disparity in *same* trials was
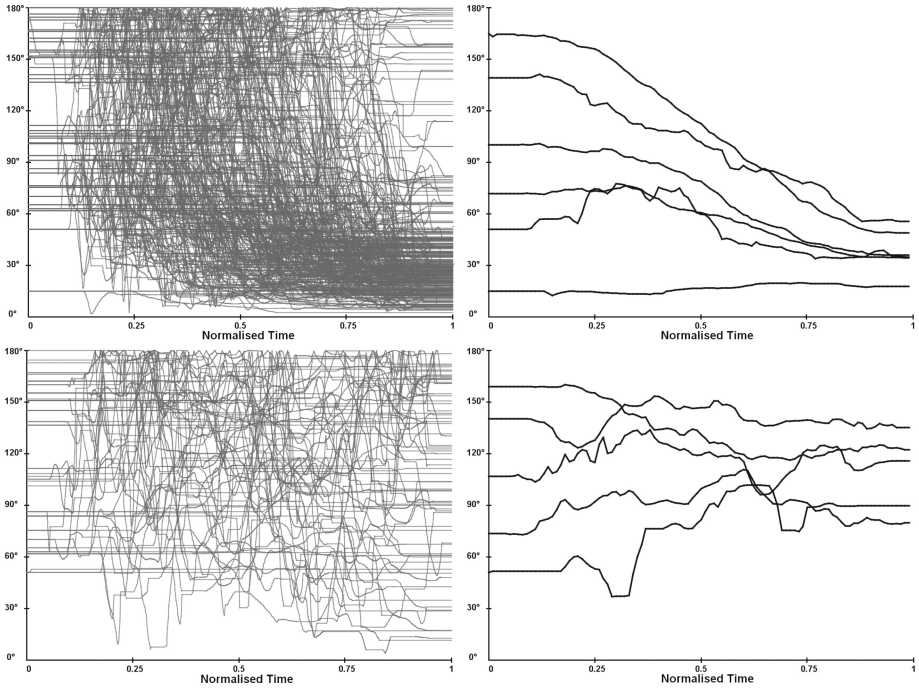
**Fig. 2.** Angular disparities for *same* tasks over time (I). Normalised time-on-task on abscissa and angular offsets (0–180°) on ordinate. *Upper row: correct, lower row: incorrect* answers. *Left column:* Observed trajectories of angular disparities over time. *Right column:* plots of average angular disparities over time for tasks grouped by 30° initial angular disparity bands.

indeed positively linearly associated with time-per-task ($p < 0.05$, $r = 0.124$). As times-per-task could not be gathered for *mental* rotation, no corresponding coefficient could be computed.

## 5   Modelling Physical Rotation

The differences in the final angular disparities between *correct* and *incorrect same* trials indicate that being able to get to final angular disparities of around 30° may be an important factor for task success. Additional comparisons revealed that *correct* trials included fewer distinct rotations, as well as more angular disparity-reducing and fewer angular disparity-increasing rotations. It thus seems quite likely that *correct* and *incorrect* trials can also be differentiated based on formal descriptions and models of the observed angular disparity changes.

It is important to note that our models are not based on data produces by students' arm or finger movements on the touch screen. Instead the models address the underlying problem solving processes. For a scheme-based modelling of the angular disparity trajectories, we concentrated on qualitative changes in
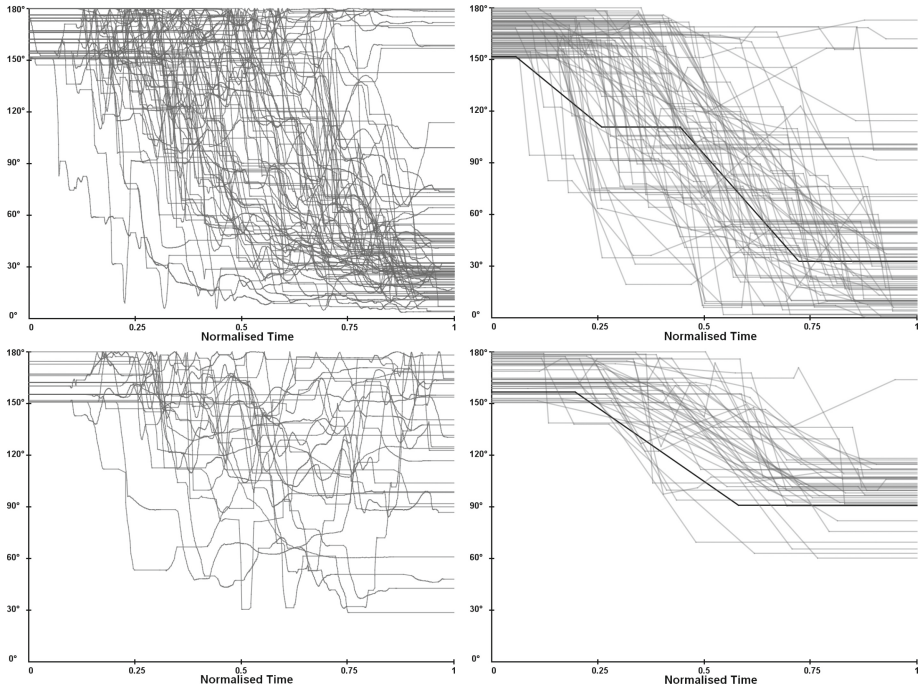
**Fig. 3.** Angular disparities for *same* tasks over time (II). For increased readability, only tasks with initial angular disparities of 150–180° are shown. *Upper row: correct, lower row: incorrect* answers. *Left column:* observed trajectories. *Right column:* synthetic trajectories generated from observed trajectories. Centroids in black, exemplifying *correct/incorrect* answer trajectories.

the trajectory, as we assume that these changes will often correspond to transitions between problem solving stages. The coding scheme was the following: each angular disparity change was coded for general direction (up, neutral, down), duration (short, medium, long), and resulting angular disparity band (one of the six 30°-bands). We then described each course as a sequence of qualitative changes, with each change represented through a tuple of values for these three variables. All courses with less than 10° accumulated rotation throughout the task (29.4% for all trials) were deemed as showing non-rotating strategies and were consequently excluded from our analysis of rotations. We assume that, for these trials, participants either employed purely analytical strategies or only employed mental rotation.

## 5.1  Generating Synthetic Trajectories

For *correct* and *incorrect same* trials, we respectively computed relative frequencies of all change-to-change transitions across all students. These relative frequencies were then used as probabilities in the construction of two Markov

models of angular disparity trajectories, one each for *correct* and *incorrect* trials. The idea was to employ the models to generate large populations of synthetic trajectories, which would be similar to the trajectories that we observed in our study. As the models are based on observed relative frequencies of change-to-change transitions, trajectories that include high frequency transitions, and are thus highly typical, could be generated most frequently. Less typical trajectories were also generated, but at suitably lower frequencies. The result was a population of synthetic trajectories that included variation while reflecting typicality. As a second advantage of our approach, we could increase the size of the set of trajectories available for subsequent steps of analysis (more synthetic trajectories than observed ones).

By analysing these large trajectory populations, prototypical trajectories could be extracted. Analysing the prototypes would then not only help us to get to the most typical rotation strategies, but hopefully also increase contrast between rotation trajectories leading to *correct* and to *incorrect* answers. In a sense, the synthetic trajectories represent minimal sets of angular disparity changes needed to describe how students typically solved tasks correctly and incorrectly.

For each of the two models, we generated 1,000 synthetic trajectories per 30° initial angular disparity band. For an illustration of these synthetic trajectories, we will largely focus on the upper-most initial angular disparity band, which includes task with initial angular disparities between 150° and 180°. Figure 3 shows these trajectories for *correct* and *incorrect* answers for *same* trials. The two graphs on the right-hand side also each include a centroid as the most typical trajectory. As we had already noted based on the observed trajectories, most of the synthetic trajectories for the correctly solved tasks show a fast decrease of angular disparity, resulting in a final angular disparity between 0° and 60°. In contrast, the decrease of angular disparity for incorrectly solved tasks is less and stops at higher final angular disparities (between 90° and 120°).

**Prototypes of Synthetic Trajectories.** To find the most typical synthetic trajectories, $k$-means clustering was applied. We computed pairwise distances between synthetic trajectories based on their qualitative descriptions (i.e., the sequences of changes) and by means of a weighted Levenshtein distance. Setting $k = 1$ produced markedly different prototypes for correct and incorrect *same* answers (shown as centroids in bold in the right column of Fig. 3). As Markov models are probabilistic, it is important to note that trajectory prototypes may vary between model runs and that the two prototypes shown in Fig. 3 are examples which may not be representative of model variability.

## 5.2   Populations of Prototypes

To address this issue, we ran each model 100 times. For the present context, we again focus on the upper-most initial angular disparity band. In each model iteration, $k$ prototypes were extracted, as before by $k$-means clustering. Figure 4
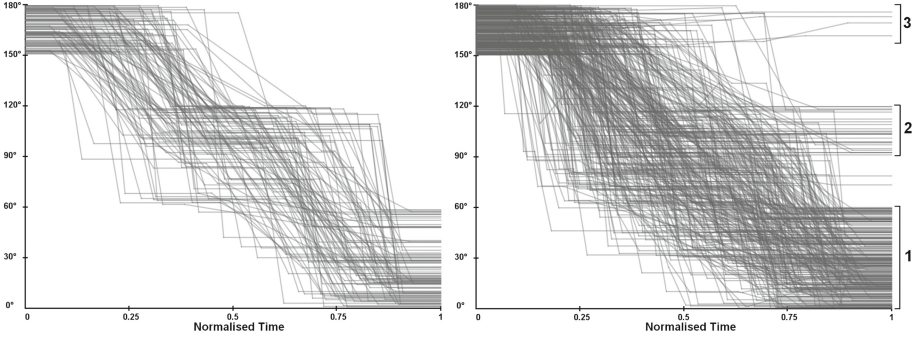
**Fig. 4.** Prototypes for *same correct* synthetic trajectories with initial angular disparities between 150° and 180°, based on 100 model runs and based on $k$-means clustering. *Left:* With ($k = 1$), *Right:* with ($k = 3$). Three distinct patterns are marked with numbers 1, 2 and 3, representing different successful rotation strategies in order of decreasing relative prominence.

shows the result of these iterations with $k = 1$ (left) and $k = 3$ (right) for *same correct* synthetic trajectories. While the left graph includes 100 prototypes (one per iteration), the right graph includes 300 prototypes (three per iteration).

For $k = 1$, only one group of typical courses can be seen, indicating the existence of one dominant rotation strategy. It involves a fast decrease of initial angular disparity until a final angular disparity between 0° and 60° is reached.

Increasing $k$ beyond 1 leads us to include also trajectories representative of secondary, tertiary, etc. strategies. For $k = 3$, three distinct groups of prototypes can be observed which differ in their relative prominence. The most dominant pattern (*strategy 1*; includes 89% of all prototypes) is the same pattern that we already observed with $k = 1$. A second pattern shows a smaller decrease of angular disparity and ends within a range of 90° to 120° (*strategy 2*; includes 9% of all prototypes). A third pattern includes prototypes with only slight changes of angular disparity, resulting in high final angular disparities between 150° and 180° (*strategy 3*; includes 2% of all prototypes). When we compare these groups of synthetic prototypes to the original trajectories that the students had produced for trials starting between 150° and 180°, we find that 72% of the original trajectories correspond to *strategy 1*, 16% to *strategy 2*, and 9% to *strategy 3*. As expected, the degree of dominance that dominant strategies possess within our population of prototypes is increased compared to their degree of dominance within the original, observed trajectories. Increasing $k$ to values above 3 did not reveal additional distinct patterns.

Table 1 shows mean values per task for all three strategies for time, accumulated distance covered during rotation (in degrees), and number of touch-based drag events during rotation. When assuming an average initial angular disparity of 165° for trials starting in the 150–180° band, *strategy 1* should on average involve a net change of angular disparity of 135° (reaching 30°, as the central value of the 0–60° band). *Strategy 2* should involve an average net change of

**Table 1.** Means per task for time, accumulated angular distance, and number of drag events during rotations for the original, observed angular disparity trajectories. The association to one of the three strategies extracted from the model is based on final angular disparity. Data is for tasks starting in the 150–180° band.

|  | Time per task | Acc. way | # of drags |
|---|---|---|---|
| Strategy 1 | 11.25 s | 227° | 12.6 |
| Strategy 2 | 10.88 s | 183° | 10.51 |
| Strategy 3 | 11.83 s | 179° | 13.53 |

60° (reaching 105°) and *strategy 3* one of 0°. The respective ratio between these values and the angular distances that were actually accumulated on average during rotation with the strategies provides us with an efficiency measure that tells us how goal-directed rotations were. For *strategy 1*, this ratio is 0.59 and for *strategy 2*, it is 0.33. For *strategy 3*, it is 0, of course, as initial and final angular disparities did not differ. In comparison, *strategy 1* is more efficient than *strategy 2*, meaning that solving a physical rotation problem with *strategy 2* involved moving through relatively more angular distance per angular distance gained in the end. Data on the number of drag events confirms that rotations with *strategy 1* were more goal-directed than with *strategy 2*, as one drag respectively covered 10.7° and 5.7° on average. Reducing angular disparity with *strategy 2* thus occurred in smaller steps than with *strategy 1*. Mean times per task were comparable between strategies.

As a last point, how strongly associated was using one of the three strategies with task success? Based on an analysis of final angular disparities, *same* trials that follow *strategy 1* were solved correctly with a probability of 96%. In contrast, only 56% of all *same* trials following *strategy 2* and 40% of all *same* trials following *strategy 3* were solved correctly.

## 6    Discussion

For the purposes of the present context, we analysed physical rotation trajectories that were captured by our app during the solving of physical rotation tasks. We were able to reproduce the angular disparity effect for *same physical* trials that were previously described by Just and Carpenter [26], and Gardony et al. [12]. This replication is important as it provides evidence for concluding that finding an angular disparity effect for 3D physical rotation tasks is likely unrelated to the choice of physical rotation control. Whereas participants in Gardony et al. [12] held a ball in hand, participants in our study used touch-based input to rotate an Arcball on an iPad display. Also, participants' age is likely not much of a factor for finding an angular disparity effect, at least between the ages of 9 and 19 (mean age in our study: 9.08 years; [12]: 19.47 years). Insofar as a general comparison of mental and physical rotation is concerned, existence of the effect in our study provides further support for assuming common or overlapping

mental processes, at least for problem solvers who are *rotators* (cf. [14]). This is useful on a methodological level, since we employ physical rotation as a proxy for and a window onto mental rotation.

We also were able to replicate the overall time courses of angular disparity for *same correct* trials that were reported by Gardony et al. [12]. Such replication again points to small or no effects of the specific physical rotation control on the involved problem solving processes.

To analyse rotation trajectories in more detail, we conducted a three-step analysis starting with a comparison of the final angular disparities between *correct* and *incorrect same* trials. We found that correctly solved *same* trials ended with a significantly lower median final angular disparity than incorrectly solved *same* trials (around 30° compared to around 120°). Achieving a final angular disparity between 0° and 60° increases the probability of correctly solving the task. With a low angular disparity, the stimuli are visually similar, which likely facilitates a visual comparison without further mental transformation.

In the second step of our analysis, we developed a qualitative description of rotation trajectories to make them more easily comparable. We coded trajectories as sequences of states, each consisting of information on direction and duration of change, as well as on final angular disparity range. Two distinct Markov models were constructed for *same correct* and *same incorrect* trials and were used to generate synthetic trajectories. A subsequent cluster analysis revealed the $k$ most typical trajectories per model run, representing the most frequently used solution strategies. As we were especially interested in successful solution strategies and the distribution of typical trajectories, we further analysed prototypical *same correct* synthetic trajectories created across 100 model runs. This analysis revealed, first, the single dominant pattern of synthetic trajectories; this represents the most frequently used successful solution strategy (*strategy 1*). We then increased $k$ step-wise to detect further distinct patterns. Using $k = 3$, we found two additional, although less frequent, successful patterns, which represent secondary and tertiary rotation strategies (*strategies 2* and *3*). Of these, *strategy 1* is the most efficient and goal-directed, based on an analysis of observed angular disparity change.

When we compare these three strategies to those reported in the literature, *strategy 1* seems to match a holistic or piecemeal mental rotation strategy where the whole object or parts thereof are mentally rotated until an angular disparity of below 25° is reached. Data on performance shows that this rotation strategy is also the most successful one. A differentiation of whether users of *strategy 1* employ holistic or piecemeal strategies does not seem possible based on an analysis of time courses of angular disparities alone. We expect that including data from additional channels (for instance, on gaze during physical rotation tasks) can help clarify on a trial-by-trial basis whether holistic or piecemeal rotation strategies are employed.

*Strategy 3* that we derived from the Markov model consists of trajectories for which initial and final angular disparities are not much different. Data on accumulated rotations shows that this does not mean that users of *strategy 3* did

not rotate. On the contrary, their rotational trajectories covered about as much overall angular distance as we observed for users of *strategy 2*. Consequently, we assume *strategy 3* to really be an example of a rotating strategy, instead of being non-rotating (e.g., analytic). We further assume that analytic strategies are to be found among those observed trajectories which we had excluded from the analysis of rotations because they hardly showed any (less than 10° accumulated throughout a trial). Performance data reveals that, compared to the other two derived strategies, *strategy 3* is least likely to lead to task success.

Interpreting *strategy 2* seems somewhat less straightforward. It involves about the same amount of overall rotation as *strategy 3*, while being directed at the final angular disparity range around 105°. *Strategy 2* is less efficient than *strategy 1* and also proceeds in smaller rotation steps. It is possible that *strategy 2* constitutes a collection of attempts at *strategy 1* that, for some reason, were not carried out to the end. Furthermore, the observation that *strategy 2*'s success rate is at chance level does suggest that terminating at higher angular disparities was unintentional. If this is so, *strategy 2* may offer potential for training-based intervention based on nudging users towards a more effective strategy.

One might ask whether some of the students' successes with the tasks were not just lucky guesses. We assume that some indeed were. However, it seems unlikely that such guesses will have had any strong influence on the process that led to extracting the three strategies. The reason for this is that there exist many ways in which a lucky guess may occur, leading to a variety of angular disparity courses. Put differently, any two lucky guesses will likely differ somewhat from one another, while any two instances of the same strategy will likely be quite similar. The consequence of lucky guessing is thus just an increase of noise surrounding the typical signatures of the three strategies.

One point still remains unclear: Did the students in our study solve tasks correctly because they achieved low final angular disparities or was successfully solving a task dependent on the process of getting there? This is a difference between an outcome- and a process-oriented view. According to Chu and Kita's view [4], it is really through the link between motor and mental rotation processes that internal computations are improved. Based on our data, we can say that trials with low final angular disparities were often solved correctly. However, we currently cannot say whether such relationship is causal or purely correlational. Assuming for the moment that it is really the process that is important for task success: Does this process need to be self-initiated to be effective, or would piloting a student to low angular disparities do the trick just as well? It seems that answering this questions will certainly be important for an interactive training of mental rotation skills. Goldin-Meadow et al. [16] compared the performance of children who were themselves gesturing during mental transformation tasks to that of children who saw someone else gesturing. They found that actually performing the gestures enhanced learning more. Whether the situation is similar with physical rotation tasks remains to be seen.

## 7   Conclusion

We presented a more detailed analysis of processes involved in solving rotation tasks than was previously available. Our method is based on time courses of angular disparity and involves qualitative descriptions of disparity changes to make the processes more easily comparable and to permit constructing generating models that spawn synthetic, but representative rotation trajectories. Based on the models that we constructed from data obtained in our study with 37 primary school students, we identified three distinct rotation-based strategies for tasks that involve high initial angular disparities. The strategies were then compared to how students rotated stimuli in our study. We expect that our general method is similarly applicable to other spatial task types. A transfer should, in particular, be easy for those types that involve sequences of mental spatial transformations, such as perspective taking or paper folding.

Our study involves two important shortcomings, which we share with other research that employed stimuli from the Vandenberg and Kuse mental rotation test: First, the initial angular disparities of the tasks were not equally distributed throughout the range of 0–180°. As can be easily seen in Fig. 2, gaps exist particularly around 120° and 30°. As a result, we focused our analysis on upper initial angular disparity bands, for which comparably many tasks exist. Secondly, *different* tasks include stimuli pairs that differ in many ways (e.g., are either mirror images or entirely different). This makes systematic cross-task comparisons of solution processes hard, if not impossible. We consequently focused on *same* tasks for the present. In an analysis of the constituent tasks of the Vandenberg and Kuse test, Geiser et al. [14] found that, for some tasks, comparisons can be easily made analytically and without mental rotation. A solution to both shortcomings lies in using different, more methodically generated stimuli sets in the future.

So far, the three strategies extracted from our *same* model have to be regarded as globally available. At this stage of analysis and with the used stimulus set, we did and could not sufficiently consider how strategy use was distributed among students (e.g., does each student have a dominant strategy?) and within each student (e.g., do individuals switch between strategies, and when?). Inter- and intra-individual differences in the use of rotation-based strategies will thus have to be addressed further in future work.

It seems important to keep in mind that time course data on stimulus rotations only shows one aspect of the problem solving process. We have already discussed that, based on our data, we could not say whether users of *strategy 1* employed mental strategies that were holistic or piecemeal in nature. Future analyses should thus include additional data for the modelling of mental and physical rotation processes, such as finger positions on the iPad or eye movements.

Last, although the Markov models were successful in increasing the contrast between high- and low-frequency rotation trajectories, the strategies that they generate are essentially local. The models are history-less insofar as they only consider probabilities of change-to-change transitions, but do not look back

(or forward) beyond each transition. In cases where problem solvers employ more global solution strategies, models should also be able to reflect more global properties of the corresponding rotation processes. Adapting our models accordingly will be among our next steps.

# References

1. Adams, D.M., Stull, A.T., Hegarty, M.: Effects of using mental and manual rotation training on mental and manual rotation performance. Spat. Cognit. Comput. **14**, 169–198 (2014)
2. Alexander, G.M., Evardone, M.: Blocks and bodies: sex differences in a novel version of the mental rotations test. Horm. Behav. **53**(1), 177–184 (2008)
3. Baenninger, M., Newcombe, N.: The role of experience in spatial test performance: a meta-analysis. Sex Roles **20**(5–6), 327–344 (1989)
4. Chu, M., Kita, S.: The nature of gestures' beneficial role in spatial problem solving. J. Exp. Psychol. Gen. **140**(1), 102–116 (2011)
5. Colangelo, N., Assouline, S.G., Gross, M.U.M.: A Nation Deceived: How Schools Hold Back America's Brightest Students, vol. 38 (2004)
6. Cooper, L.A., Shepard, R.N.: Chronometric studies of the rotation of mental images. In: Visual Information Processing, pp. 15–20 (1973)
7. Eisenegger, C., Herwig, U., Jäncke, L.: The involvement of primary motor cortex in mental rotation revealed by transcranial magnetic stimulation. Eur. J. Neurosci. **25**(4), 1240–1244 (2007)
8. Feng, J., Spence, I., Pratt, J.: Playing an action video game reduces gender differences in spatial cognition. Psychol. Sci. **18**(10), 850–855 (2007)
9. Frick, A., Wang, S.H.: Mental spatial transformations in 14- and 16-month-old infants: effects of action and observational experience. Child Dev. **85**(1), 278–293 (2014)
10. Ganis, G., Keenan, J.P., Kosslyn, S.M., Pascual-Leone, A.: Transcranial magnetic stimulation of primary motor cortex affects mental rotation. Cereb. Cortex **10**(2), 175–180 (2000)
11. Gardner, H.: Frames of Mind: The Theory of Multiple Intelligences. Basic Books, New York (2011)
12. Gardony, A.L., Taylor, H.A., Brunyé, T.T.: What does physical rotation reveal about mental rotation? Psychol. Sci. **25**(2), 605–612 (2014)
13. Gardony, A.L., Taylor, H.A., Brunyé, T.T., Wolford, G.L.: To rotate or not to rotate: strategy differences in mental rotation. In: Proceedings of the 55th Annual Meeting of the Psychonomic Society, Long Beach, CA, USA (2014)
14. Geiser, C., Lehmann, W., Eid, M.: Separating "rotators" from "nonrotators" in the mental rotations test: a multigroup latent class analysis. Multivar. Behav. Res. **41**(3), 261–293 (2006)
15. Gersmehl, P.J., Gersmehl, C.A.: Spatial thinking by young children: neurologic evidence for early development and "educability". J. Geogr. **106**(5), 181–191 (2007)
16. Goldin-Meadow, S., Levine, S.C., Zinchenko, E., Yip, T.K.Y., Hemani, N., Factor, L.: Doing gesture promotes learning a mental transformation task better than seeing gesture. Dev. Sci. **15**(6), 876–884 (2012)
17. Hamrick, J.B., Griffiths, T.L.: What to simulate? inferring the right direction for mental rotation. In: Annual Conference of the Cognitive Science Society, pp. 577–582 (2014)

18. Heckman, J.J.: Skill formation and the economics of investing in disadvantaged children. Science **312**(5782), 1900–1902 (2006)
19. Heil, M., Jansen-Osmann, P.: Sex differences in mental rotation with polygons of different complexity: do men utilize holistic processes whereas women prefer piecemeal ones? Q. J. Exp. Psychol. **61**(5), 683–689 (2008)
20. Heil, M., Rösler, F., Link, M., Bajric, J.: What is improved if a mental rotation task is repeated-the efficiency of memory access, or the speed of a transformation routine? Psychol. Res. **61**(2), 99–106 (1998)
21. Huttenlocher, J., Newcombe, N., Vasilyeva, M.: Spatial scaling in young children. Psychol. Sci. **10**(5), 393–398 (1999)
22. Jansen, P., Heil, M.: Suitable stimuli to obtain (no) gender differences in the speed of cognitive processes involved in mental rotation. Brain Cognit. **64**(3), 217–227 (2007)
23. Jansen, P., Heil, M.: The relation between motor development and mental rotation ability in 5-to 6-year-old children. Int. J. Dev. Sci. **4**(1), 67–75 (2010)
24. Jansen, P., Schmelter, A., Kasten, L., Heil, M.: Impaired mental rotation performance in overweight children. Appetite **56**(3), 766–769 (2011)
25. Jansen, P., Schmelter, A., Quaiser-Pohl, C., Neuburger, S., Heil, M.: Mental rotation performance in primary school age children: are there gender differences in chronometric tests? Cognit. Dev. **28**(1), 51–62 (2013)
26. Just, M.A., Carpenter, P.A.: Eye fixations and cognitive processes. Cognit. Psychol. **8**(4), 441–480 (1976)
27. Kelnberger, M.: Fit für den Übertritt - Bildungsstandards Mathematik: Tests zur Leistungsfeststellung (2008)
28. Khooshabeh, P., Hegarty, M., Shipley, T.F.: Individual differences in mental rotation: piecemeal versus holistic processing. Exp. Psychol. **60**(3), 164 (2012)
29. Kosslyn, S.M., Thompson, W.L., Wraga, M., Alpert, N.M.: Imagining rotation by endogenous versus exogenous forces: distinct neural mechanisms. NeuroReport **12**(11), 2519–2525 (2001)
30. Kozhevnikov, M., Hegarty, M.: A dissociation between object manipulation spatial ability and spatial orientation ability. Mem. Cognit. **29**(5), 745–756 (2001)
31. Linn, M.C., Petersen, A.C.: Emergence and characterization of sex differences in spatial ability: a meta-analysis. Child Dev. **56**(6), 1479–1498 (1985)
32. Logie, R.H., Pernet, C.R., Buonocore, A., Sala, S.D.: Low and high imagers activate networks differentially in mental rotation. Neuropsychologia **49**(11), 3071–3077 (2011)
33. Marmor, G.S.: Development of kinetic images: when does the child first represent movement in mental images? Cognit. Psychol. **7**, 548–559 (1975)
34. Metzler, J., Shepard, R.N.: Transformational studies of the internal representation of three-dimensional objects. In: Solso, R.L. (ed.) Theories of Cognitive Psychology: The Loyola Symposium. Lawrence Erlbaum, Potomac (1974)
35. Newcombe, N.S., Frick, A.: Early education for spatial intelligence: why, what, and how? Mind Brain Educ. **4**(3), 102–111 (2010)
36. Peters, M., Battista, C.: Applications of mental rotation figures of the Shepard and Metzler type and description of a mental rotation stimulus library. Brain Cognit. **66**(3), 260–264 (2008)
37. Quaiser-Pohl, C., Neuburger, S., Heil, M., Jansen, P., Schmelter, A.: Is the male advantage in mental-rotation performance task independent? on the usability of chronometric tests and paper-and-pencil tests in children. Int. J. Test. **14**(2), 122–142 (2014)

38. Shepard, R.N., Metzler, J.: Mental rotation of three-dimensional objects. Science **171**, 701–703 (1971)
39. Shoemake, K.: Arcball: a user interface for specifying three-dimensional orientation using a mouse. Graph. Interface **92**, 151–156 (1992)
40. Terlecki, M.S., Newcombe, N.S., Little, M.: Durable and generalized effects of spatial experience on mental rotation: gender differences in growth patterns. Appl. Cognit. Psychol. **22**(7), 996–1013 (2008)
41. Thüringer Ministerium für Bildung, Wissenschaft und Kultur: Lehrplan für die Grundschule und für die Förderschule mit dem Bildungsgang Grundschule - Mathematik (2010)
42. Tzuriel, D., Egozi, G.: Gender differences in spatial ability of young children: the effects of training and processing strategies. Child Dev. **81**(5), 1417–1430 (2010)
43. Uttal, D.H., Cohen, C.A.: Spatial thinking and STEM education: when, why and how? Psychol. Learn. Motiv. **57**(2), 147–181 (2012)
44. Uttal, D.H., Meadow, N.G., Tipton, E., Hand, L.L., Alden, A.R., Warren, C., Newcombe, N.S.: The malleability of spatial skills: a meta-analysis of training studies. Psychol. Bull. **139**(2), 352 (2013)
45. Vandenberg, S.G., Kuse, A.R.: Mental rotations, a group test of three-dimensional spatial visualization. Percept. Motor Skills **47**(2), 599–604 (1978)
46. Wai, J., Lubinski, D., Benbow, C.P.: Spatial ability for stem domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. J. Educ. Psychol. **101**(4), 817 (2009)
47. Wexler, M., Kosslyn, S.M., Berthoz, A.: Motor processes in mental rotation. Cognition **68**(1), 77–94 (1998)
48. Wiedenbauer, G., Jansen-Osmann, P.: Manual training of mental rotation in children. Learn. Instr. **18**(1), 30–41 (2008)
49. Wohlschläger, A., Wohlschläger, A.: Mental and manual rotation. J. Exp. Psychol. Hum. Percept. Perform. **24**(2), 397–412 (1998)
50. Wright, R., Thompson, W.L., Ganis, G., Newcombe, N.S., Kosslyn, S.M.: Training generalized spatial skills. Psychon. Bull. Rev. **15**(4), 763–771 (2008)
51. Zander, S., Wetzel, S., Bertel, S.: Rotate it! - effects of touch-based gestures on elementary school students' solving of mental rotation tasks. Comput. Educ. **103**, 158–169 (2016)