**A novel approach to analyzing all trials in chronometric mental rotation
and description of a flexible extended library of stimuli**

Leonardo Jost and Petra Jansen

University of Regensburg

Author Note

Leonardo Jost, Faculty of Human Sciences, University of Regensburg;

Petra Jansen, Faculty of Human Sciences, University of Regensburg.

Correspondence concerning this article should be addressed to Leonardo Jost, Faculty of
Human Sciences, University of Regensburg, Universitätsstraße 31, 93053 Regensburg,
Germany.

Email: Leonardo.jost@ur.de ORC ID: 0000-0003-3508-4315

This is an Accepted Manuscript of an article published by Taylor & Francis in Spatial
Cognition & Computation: An Interdisciplinary Journal on 03 May 2020, available online:
<https://www.tandfonline.com/doi/full/10.1080/13875868.2020.1754833>

17 **A novel approach to analyzing all trials in chronometric mental rotation** 18 **and description of a flexible extended library of stimuli**

19 Mirrored stimuli in chronometric mental rotation tests cannot be rotated into
20 congruence and are typically discarded from analysis. We present a novel design using
21 three figures such that congruence is achieved for all trials. For design construction we
22 consider the geometrical properties of rotation and mirroring.

23 We have experimentally tested the approach with forty-one German sport students.
24 Results analyzed with linear mixed modeling show small differences but are
25 comparable to the classical two-figure design. Based on our findings we suggest our
26 new layout for future chronometric mental rotation. This will improve analysis of
27 accuracies while increasing power.

28 Moreover, we present a database of cube figures and code for flexible generation of
29 stimuli, which we make available on GitHub at
30 <https://github.com/LeonardoJost/MRlibrary>.

31 *Keywords:* mental rotation; stimuli; library; test

Introduction

Spatial abilities are important in everyday activities and they are related to various abilities like mathematical ability (Xie, Zhang, Chen, & Xin, 2019) or problem solving (Geary, Saults, Liu, & Hoard, 2000). An important spatial ability, especially an intrinsic-dynamic ability according to the classification of Uttal et al. (2013), is mental rotation, the skill to imagine rotated objects in mind. An increased mental rotation performance is related to the attendance of a high school with many spatial school subjects such as chemistry, design, or biology (Moè, 2016) or a STEM degree (Hausmann, 2014). Due to the significance of the mental rotation performance it seems to be necessary to design well-defined mental rotation test to investigate this specific spatial performance. Because the ability to mentally rotate items is not only of theoretical scientific interest but relates for example to STEM education (Moè, Jansen, & Pietsch, 2018), it is the main goal of this study to describe a new experimental mental rotation design, with which all trials can be analyzed.

To investigate the mental rotation ability, psychometric (paper-pencil) and chronometric (computer based) tasks exist. Until now, in classical chronometric mental rotation tests (Shepard & Metzler, 1971), participants are presented two figures and they have to mentally rotate one figure to decide if the figures are the “same”, which means they can be rotated into congruence, or “different”, which means mirror-reversed. This decision of non-mirrored or mirror-reversed object requires an object-based transformation (Voyer, Jansen, & Kaltner, 2017). In this kind of transformation the observer's position remains fixed and the object is moved mentally in relation to the surrounding environment. In the other class of mental transformations, egocentric transformations, figures, which typically are human bodies or body parts, are processed by a transformation of one's own perspective (Zacks & Tversky, 2005). However, mental rotation can be dissociated from the ability to reorient the imagined self, which is mostly measured by perspective-taking tests (Kozhevnikov &

Hegarty, 2001). A highly discussed issue in mental rotation research is the existence of gender differences. However, there is an ongoing debate as to whether and how gender influences chronometric mental rotation tests (Jansen-Osmann & Heil, 2007).

Next to influences on performance of stimulus properties such as dimensionality (Jolicœur, Regehr, Smith, & Smith, 1985; Shepard & Metzler, 1988), the role of distractors (Boone & Hegarty, 2017) and interpreting results regarding the usage of different strategies (Khooshabeh & Hegarty, 2010) to date, there is one main methodological concern in the use of chronometric mental rotation tests: Because mirrored figures cannot be brought into congruence those items have been deleted mostly from analysis so far (Jolicœur et al., 1985, Shepard & Metzler, 1971). This incongruence is rarely further analyzed but is reflected in reaction time differences of almost one second in the analysis of Shepard and Metzler (1971) and around 500ms for cube figures in Voyer and Jansen (2016). As mirrored figures typically contribute about half of the tasks, disregarding them means losing a noticeable amount of power or testing time in experiments (Brysbaert & Stevens, 2018). Moreover, analyzing accuracy only on non-mirrored stimuli disregards both false alarms and correct rejects, which might show a different pattern than hits and misses. Analyzing sensitivity instead of accuracy incorporates false alarms (Stanislaw, & Todorov, 1999). However, as false alarms occur only on mirrored stimuli, they cannot be matched to an angle of rotation and only global analyses of sensitivity are appropriate.

To overcome the shortcoming of the former chronometrical mental rotation tests we present a novel approach, in which participants are shown three figures instead of two. A similar approach has already been used by Wohlschläger and Wohlschläger (1998) and by two subsequent studies from Wohlschläger (2001) and Krüger and Krist (2009) while investigating the common processing of manual and mental rotation. However, because they were not interested in the development of the design, they did not differ in their analysis

between mirrored and non-mirrored items. They found that reaction time averaged across their mirrored and non-mirrored stimuli demonstrated an increase with increasing angular disparity. In the design presented here we will call two figures the base figures, which are mirrored to each other, and the third the stimulus figure, which is a rotated version of one of the base figures. Instead of a same/different decision, participants are tasked to find out if the stimulus can be rotated into congruence with one or the other base figure. As the base figures are mirrored to each other and not mirror-symmetrical, the stimulus is congruent to one and only one base figure. Thus, the rotation into congruence is independent of stimulus orientation.

Goal of the study

As mentioned above, the main goal of this study is to evaluate the new proposed mental rotation paradigm. While in theory the task at hand can be solved by ignoring one base figure and proceeding as in the case of two figures, we believe that presentation of a unique 0°-condition for all stimuli will show the same behavior for both same and different stimuli (or matching the left or the right base figure in this study). Moreover, we believe that for all stimuli the same mental transformations as in the two-figure case are performed and we expect to find the well-known mental rotation effects: An increase of reaction time and a decrease of accuracy with increased angles of rotation (Shepard & Metzler, 1971). Furthermore, we predict that participants improve over time with larger improvements for larger degrees. By design, we expect these effects not to differ by the side of the correct answer, although we presume a possible effect due to the left/right arrangement in accordance with the reading direction of our participants. Because of the geometric considerations of the rotation axes presented in the method section, we hypothesize to find no differences of the aforementioned effects between rotation axes. As a secondary hypothesis, we expected to

find better performances for participants with previous experience in chronometric mental rotation experiments. At least, short-term benefits for mental rotation training on mental rotation tasks have been shown so far (Meneghetti, Cardillo, Marmarella, Caviola, & Borella, 2017). Due to the debate on gender differences in chronometric mental rotation tests (e.g. Jansen-Osmann & Heil, 2007) gender was not regarded as an additional factor.

In addition to the empirical evaluation of the test, we have created an extended, flexible library of stimulus material and describe its properties and utilization.

Method

Participants

A total of 41 German students (32 females, 9 males) participated in this study due to study credit. On average, participants were 21.3 years old ($SD=1.8$). In our sample, we included participants with self-reported previous experience in mental rotation experiments employing the two-figure design (24 with experience, 17 without). For appropriate power, Brysbaert and Stevens (2018) recommend at least 1600 observations per condition independent of effect sizes of other work. As we expected to exceed an average of 40 stimuli per participant per condition in our design, this should yield high power for the main hypotheses. Supporting this we found 14447 usable measurements in the smallest dataset equivalent to 1805 observations per condition for the triple interactions of interest. For the secondary hypotheses, Meneghetti et al (2017) report large effect sizes of $d>1$ for transfer within chronometric mental rotation. G*power (Faul, Erdfelder, Lang, & Buchner, 2007) shows a power of .80 for a between subjects t-test at an effect size of $d=0.8$.

Stimulus library

Due to the work of Peters and Battista (2008) a stimulus library of cube figures is already

available. In an attempt to expand the applications and research possibilities, we have written an extended, more flexible generation process in R (R Core Team, 2018). To keep comparisons with older work valid, we have used the same cube structure for the base figures as Peters and Battista (2008), although we do not exactly reproduce the rotated versions. In the supplementary material, we describe the reasoned small changes we have made as well as possible extensions and applications. Both pregenerated cube figures and the code for generation of individual figures are available at <https://github.com/LeonardoJost/MRlibrary>. Upon usage, we advise to state all values of the parameters given in table 1.

Rationale for figure layout

Different layouts are possible to achieve the proposed mental rotation task with three figures with some possibilities presented in figure 1 and many more possible layouts not depicted. We have decided to present the three figures arranged as in figure 1a and 4 for three following main reasons: First, the usage of screen space allows larger images while avoiding overlapping. Second, the layout provides congruence between hand position, response key layout, and figure placement (left/right). Third, the optical distances between stimuli and base figures are small and comparable, while also keeping discrimination between base and stimulus figures clear for participants.

Please insert figure 1 about here

Geometric considerations of axes and mirroring planes in three dimensions

The necessity for the development of the new test presented here is due to the incongruence of mirrored stimuli. In this section, we will review this incongruence and describe geometrical properties of mirroring planes and rotation axes, which we have used as the basis for our design. For the most part, we only consider the natural planes and axes for mirroring and rotation, i.e. the xy-, xz-, and yz-plane and the x-, y-, and z-axis. For the naming of the

axes, we have used the convention that the vertical axis is the z-axis with larger values for higher points, the horizontal axis is the x-axis with larger values for points further to the right and the depth axis is the y-axis with larger values for points further back. If these axes pass through the center of an object with the front pointing towards the observer, the resulting spins are known as pitch for the rotation around the x-axis, roll for the rotation around the y-axis, and yaw for the rotation around the z-axis. It also has to be noted, that mirrored stimuli are images of the mirrored figure. As base rotation angles are employed in the 0°-position, these are not actual mirror images. This is demonstrated in figure 2, where images of mirrored figures are shown. While the three-dimensional figures are mirrored to each other, the two-dimensional images of them are not.

Incongruence of mirrored stimuli

While a mirrored figure is clearly unique in shape, depending on the orientation of the mirroring plane, the mirror image appears at different rotated positions. In figure 2, the images resulting from a mirroring across the xy-plane and the yz-plane differ by a rotation around the y-axis by 180°. More generally, two mirror figures deviate by twice the angle at which the two mirroring planes are rotated to each other. As a result, there exists no unique 0°-position and thus no angular disparity for mirrored stimuli. This problem also occurs in two dimensions, as rotation of mirroring lines also produces rotated stimuli. A case, in which this incongruence might be lessened are for example pictures of human or animal figures, where a clear upright position exists, or mental rotation tasks with egocentric transformations. Rotations around the z-axis, which maintain the upright position, might still cause problems, although one might argue in favor of a unique front facing position, which would also solve this incongruence.

Please insert figure 2 and 3 about here

177 *Difference and relationships between mirroring planes and rotation axes*

178 As we choose to present two mirrored base figures and by the incongruence no unique way of
179 choosing the mirrored figure exists, the question of choosing a base orientation arises. We
180 propose to use the mirroring plane that optically lies between the two base figures, such that
181 they can more easily be recognized as mirrored shapes of each other. As a result, we use the
182 yz-plane for the mirroring in our layout, while e.g. in the layout of figure 1c we would use the
183 xy-plane. As we cannot depict a layout in depth on a two dimensional screen, we ignore
184 mirroring across the xz-plane in this step.

185 Moreover, mirroring switches the image along one axis while rotation keeps one axis
186 fixed. Mirroring across one plane switches the values along the axis that is perpendicular to
187 that plane, while keeping values constant along the axes that span the plane. Rotation on the
188 other hand keeps the coordinates along the rotation axis constant, while the other two change.
189 As can be seen in figure 2, mirroring across the yz-plane switches the coordinates along the
190 x-axis, while keeping the others the same, i.e. left and right are reversed. The rotation can be
191 seen in figure 3, where, as the figure is rotated around the x-axis, the left and right parts do
192 not change, but front/back and top/bottom are moved by the rotation. As a result, if a
193 mirroring across the yz-plane and a rotation around the x-axis were used and the rotation axis
194 were known, no rotation strategy would be necessary to solve the task, as a simple left/right
195 comparison would be sufficient. Thus, it seems sensible to choose the rotation axes from the
196 axes that span the mirroring plane. In our case of the yz-plane these are the y- and the z-axis,
197 as we restrict ourselves to the natural axes.

198 Another problem with a combination of mirroring and rotation is demonstrated in
199 figure 4. A rotation by 180° around the y-axis of the left base figure can easily identified as a
200 mirroring of the right base figure across the xy-plane and similarly a rotation around the z-
201 axis as a mirroring across the xz-plane. Because participants will be informed that only one

side can be correct and that the two base figures are mirrored to each other, the task can be solved by exclusion. While this strategy is demonstrated in the case of a rotation by 180°, it is theoretically applicable to all rotation angles, but reaches the natural mirroring planes only at 0° and 180° (and for a rotation of 0° a much simpler solution is available as two figures are identical). As generally larger reaction times are found in the case of mirrored stimuli in classical mental rotation (Shepard & Metzler, 1971; Voyer & Jansen, 2016), where the same strategy is applicable, this strategy seems too complex to have a major impact on results.

On a further note, the three axes differ in dimension, as rotation around the y-axis is only a rotation in the picture plane, while the x- and z-axis produce rotations in depth, but typically no systematic differences between in-depth rotation and picture-plane rotation are found (Shepard & Metzler, 1971).

Please insert figure 4 about here

Material

Mental Rotation

Stimulus presentation and response handling were controlled with Presentation® software (Version 20.1 Build 12.04.17, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com) on a Dell Latitude E7240 Laptop (12.5" screen, 1366x768, 60Hz).

Participants were seated in front of a desk with the laptop on it and were free to adjust both the laptop and their seating position.

The stimuli were generated from the stimulus library with the parameters given in table 1. Stimulus images were then resized to 400px times 400px and presented according to the layout in figure 1a and 4 until a response was given. The vertical shift of the images was set to 150 pixels (above the center for the base figures and below the center for the stimulus figure) and the horizontal shift to 300 pixels (left and right of the center for the base figures).

Please insert table 1 about here

Between tasks in the practice session, participants received feedback for 1000ms (✓-right, ✗- wrong) shown at the center of the screen at font size 40. In the main session, participants received no feedback and a fixation cross (“+”) was shown at the center of the screen for 500ms.

Stimuli were presented until 10 minutes (practice session) or 30 minutes (main session) passed, such that all stimuli were shown once in random order, generated by shuffling in Presentation® software. Afterwards, every stimulus was shown again twice in another random order. Between the first two occurrences of the same stimulus, at least 20 other stimuli were shown. The orientation of the left and right base figure was randomized.

Participants were digitally instructed to press the left mouse button on the touchpad of the laptop with their left index finger if the stimulus could be rotated into congruence with the left base figure. If the stimulus could be rotated into congruence with the right base figure, participants should press the right mouse button with their right index finger. Participants were asked to answer as quickly and as precisely as possible.

Reaction time, accuracy, stimulus type (model, rotation angle, rotation axis, stimulus orientation, base orientation), and time since start of the session were recorded.

Demographics

A digital questionnaire was used to collect demographic information. Participants were asked about their previous experience with mental rotation (participants had to indicate if they had or had not participated in other mental rotation experiments before), age (in years), sex/gender (male, female, or other), information about their menstrual cycle, physical and musical activity, and handedness.

249 *Procedure*

250 Participants completed a practice session of 10 minutes and continued with the main session
251 of 30 minutes after a self paced break. Following the main session, participants answered a
252 digital questionnaire regarding demographic data. Total durations of individual sessions were
253 known to participants before the start of the experiment and were shown again before each
254 session. For the duration of the experiment and the questionnaire, participants were alone in
255 the experimental room.

256 *Statistical analysis*

257 The accuracy and time of response of each trial were used as dependent variables and the
258 angular disparity, time (since start of the session), the side of the correct answer, the rotation
259 axis, and previous experience of participants were used as independent variables.

260 Outliers were determined by a deviance of more than three standard deviations from
261 the mean reaction time of all stimulus pairs with the same rotation angle and were excluded
262 from all analyses and reaction time was additionally only analyzed on correct responses.
263 Moreover, as the rotation axis is not well-defined for the 0°-condition, all analyses including
264 axes were calculated only on the rotated stimuli. For all effects not containing axes, the
265 analysis was calculated twice, once on the dataset including the non-rotated stimuli and once
266 on the dataset of only the rotated stimuli, possibly including effects of axes as moderating
267 effects. By this procedure, 302 of 17900 trials were deemed as outliers. Of the remaining
268 trials, 2073 were incorrect responses and 1142 trials used no rotation.

269 Statistical analysis was performed according to Jost, Weishäupl, and Jansen (2019)
270 with linear mixed models using lme4 package (version 1.1-21; Bates, Maechler, Bolker, &
271 Walker, 2015) in R (version 3.5.1; R Core Team, 2018). Model parameters were estimated by
272 maximum likelihood estimation using bobyqa algorithm wrapped by optimx package (version

2018-7.10; Nash & Varadhan, 2011) as optimizer. Model fit was calculated by using likelihood ratio tests to compare models with and without the fixed effect of interest. The resulting p-values were compared to a significance level of .05. For multiple comparisons of the same variables, the significance level was Bonferroni corrected. Visual inspection of residual plots did not reveal deviations from homoscedasticity or normality in any model.

Where possible, we report both the unstandardized effect sizes and confidence intervals calculated by using parametric bootstrapping with 1000 simulations in line with recommendations of Baguley (2009) and Pek and Flora (2018). While standardized effect sizes are routinely used for power analysis and meta analyses, unfortunately there does not exist an agreed upon way to compute standardized effect sizes in linear mixed models (Feingold, 2009; Hedges, 2007; Rights & Sterba, 2019). Nevertheless, linear mixed models offer several advantages over traditional use of ANOVAs. For example, linear mixed models allow simultaneous analysis of by-participant and by-item variances and thus eliminating the need to average over participants or items, while also allowing analysis of unbalanced data and achieving higher statistical power (Barr et al., 2013, Hilbert et al., 2019). These allowed the experiment to be controlled by time with participants finishing different numbers of trials and enabled the analysis of time as an independent variable.

Model building was based on the research of Barr et al. (2013) and Bates, Kliegl, Vasishth, and Baayen (2015), starting with a model with random intercepts and slopes for every appropriate fixed effect and reducing the model complexity by dropping non-significant variance components. Non-significant fixed effects were further stepwise removed from the model, such that effects which least decreased model fit were removed first and a model containing only significant fixed effects remained. Non-significant effects were then tested for an improvement of model fit by inclusion in the resulting model, while significant effects were tested for worsening of model fit by exclusion of the effect. Main effects for

significant interactions were tested separately by splitting the interaction. The resulting models for each parameter are described in the results section.

As there is ongoing discussion about model selection based on theory or data or preferring complex instead of simpler models, we expect future research to cast doubt on the optimality of the currently suggested models. Although models without random slopes seem too anti-conservative (Barr et al., 2013) we report results of these simplest models for easier comparison.

Results

Descriptive statistics

Means and standard deviations for both reaction time and accuracy are provided in table 2 for both side and degree as the main effects in question. Due to the time-controlled nature of the experiment, participants performed different number of trials. Average reaction time and accuracy are calculated for every participant. Mean and standard deviations are then reported between participants. Behavioral data for other conditions and summarized demographic data can be found at <https://github.com/LeonardoJost/MRExperiment>.

Please insert table 2 about here

Reaction time

For the analysis of reaction time on the dataset of only rotated stimuli, model building resulted in a model with random intercepts and random slopes for degree and time (since start of the session) by participant and random intercepts by model. The interactions degree*time*side and degree*axis*side as well as the included two-way interactions and main effects, and the effects of previous experience were analyzed as fixed effects. Significant effects were found for experience, degree*time, degree*side, degree*axis, degree,

and time. Reaction time improved with experience and time and increased with angular disparity. The effect of angular disparity was lower with increased time, for answers on the right side, and for rotations around the z-axis. (See table 3)

Please insert table 3 about here

In a model without random slopes, the effects of experience ($\chi^2(1)=0.13$, $p=.719$) and degree*side ($\chi^2(1)=3.53$, $p=.060$) were not significant, but otherwise no differences in significance were found.

For the analysis of reaction time on the dataset of all stimuli, model building resulted in a model with random intercepts and random slopes for degree and time by participant and random intercepts by model. The interaction degree*time*side and the effects of previous experience were analyzed as fixed effects. Significant effects were found for experience, degree*time, degree*side, degree, time, and side. Compared to the model for only the rotated stimuli, reaction times were significantly slower for answers on the right side, but otherwise no differences in significance were found. (See table 4)

Please insert table 4 about here

In a model without random slopes, the effects of experience ($\chi^2(1)=0.11$, $p=.740$) and side ($\chi^2(1)=3.65$, $p=.056$) were not significant, but otherwise no differences in significance were found.

Accuracy

Accuracy was analyzed by a general linear mixed model, which used a binomial distribution. For the analysis on the dataset of only rotated stimuli, model building resulted in a model with random intercepts and random slopes for degree, time (since start of the session), and axis by participant and random intercepts and random slopes for time by model. The interactions degree*time*side and degree*axis*side and the effects of previous experience

were analyzed as fixed effects. Significant effects were found for degree*axis, degree, and time. Accuracy was significantly higher with increased time and significantly lower with increased angular disparity. The effect of degree was lower for rotations around the z-axis. (See table 5)

Please insert table 5 about here

In a model without random slopes, the effects of degree*time ($\chi^2(1)=5.12$, $p=.024$), degree*side ($\chi^2(1)=3.94$, $p=.047$), and axis ($\chi^2(1)=7.41$, $p=.006$) were significant, but otherwise no differences in significance were found.

For the analysis of accuracy on the dataset of all stimuli, model building resulted in a model with random intercepts and random slopes for degree and time by participant and random intercepts and random slopes for time by model. The interaction degree*time*side and the effects of previous experience were analyzed as fixed effects. Significant effects were found for degree*side, degree, and time. Compared to the model for only the rotated stimuli the decrease of accuracy with increasing angular disparity was significantly smaller for answers on the right side, but otherwise no differences in significance were found. (See table 6)

Please insert table 6 about here

In a model without random slopes, the effect of degree*time ($\chi^2(1)=7.39$, $p=.007$) was significant and the effect of time ($\chi^2(1)=3.58$, $p=.058$) was not significant, but otherwise no differences in significance were found.

As we observed nonlinear behavior in the accuracy data averaged by participant, we performed separate comparisons by degree for both rotation axis and the correct side. For this, we used a model with random intercepts and random slopes for time by both participant and model. As fixed effects, we included both axis and side for all angles for comparison purposes. At 0° answers on the right side were significantly less accurate than answers on the

left side, while at 180° they were significantly more accurate. No differences were found at 45°, 90°, and 135°. Rotations around the z-axis produced significantly more accurate answers at 135° and 180° compared to rotations around the y-axis. (See table 7)

Please insert table 7 about here

Linearity of data

As we observed nonlinear behavior in the data, we analyzed the influence of additional polynomial functions of angular disparity on both reaction time and accuracy. In these cases, a significant improvement of model fit was found for quadratic and cubic polynomials. All estimated polynomials were monotonously increasing for reaction time and monotonously decreasing for accuracy in the value range of interest. (See table 8)

Please insert table 8 about here

Discussion

Our study provides insight that the newly developed test is applicable for measuring mental rotation ability. With this test all stimuli can be analyzed, although small differences between the sides of the correct base figure persist. We have found some minor differences in reaction time but some larger differences in accuracy. While the behavior of reaction time and accuracy by degree shows changes by side and with different rotation axes, the overall results are comparable to the behavioral data of “same” stimuli in the design of Shepard and Metzler (1971) with two images regarding effects of degree. That is, for both sides of the congruent base figure we observe the behavior expected of “same” stimuli. Moreover, if only one base figure were used to solve the task, we should have found reaction time differences comparable to those between “same” and “different” trials in two-figure mental rotation tasks. As this is not the case, we assume that the unique 0°-position is used by participants and the same mental rotation process is measured in all trials.

Regarding reaction time, we have found the expected relationship to degree and time. Reaction time increases with larger angles and decreases over time, with larger improvements at larger angles. We have found differences in the interaction of degree and side, but on both sides our results show the expected increase of reaction time at larger angles. Moreover, the differences by side are smaller than by axis, which are typically assumed to be the same. Possibly of the same magnitude are effects of rotation direction, which were not analyzed and are also typically grouped together. Additionally, the improvement over time at all angles did not differ by side, which leads us to the assumption, that the same mental transformation processes are used on both sides. This is also supported by the lower reaction times of participants with previous experience with the two-figure layout, although learning effects are not fully understood (Heil, Rösler, Link, & Bajric, 1998; Meneghetti et al., 2017; Rahe, Ruthsatz, Jansen, & Quaiser-Pohl, 2019; Uttal et al., 2013).

For accuracy we have found differences by side for angular disparities of 0° and 180° as well as differences by axis. Compared to reaction time, accuracy shows no significant differences for previous experience and for the interaction of degree and time, but results point in the same direction as for reaction time. In the case of side and axis, the results indicate higher accuracies at lower reaction times and cannot be justified by speed-accuracy tradeoffs. The differences in accuracy at 0° are quite surprising, as no mental rotation is necessary, yet average accuracy was only at 91% for answers on the right side (compared to 97% for the left side). One possible explanation might be the left to right reading direction and thus figures were possibly also compared in that order, although reaction time differences would be better explained by this than accuracy differences. Also, we have observed a quite unexpected increase in accuracy for rotations around the z-axis at larger angles. While this might be influenced by the layout of the figures or the base angles of the figures, an interaction with the side of answer was not significant. A possible influence can be seen in

figure 4 for one model, but the other models show a similar position of the characteristic arm. For a rotation around the z-axis, the characteristic arm of the figure stays close to the two base figures, whereas for a rotation around the y-axis the optical distance between stimulus and base figures increases. By the same matter, there is a possible overlap of the figure with the fixation cross between trials for rotations around the z-axis. Another difference can be seen in the base angles, which produce different directions of the front- and top-facing arms, which would point straight forward or upwards if no base angles were used. As these explanations are speculative, it remains unclear, how and why accuracy differences by axis occur. Further research needs to be done on reproducibility and explanations of these phenomena and accuracy should be interpreted with more caution.

Similarly as for reaction time, overall accuracy patterns show smaller differences by side than by axis and no interactions with time, which can be interpreted in support of the notion that the same mental transformations are performed on both sides as in the two-figure case.

Average reaction times are also similar to other studies using comparable stimuli but different designs. For example, Jost, Weishäupl, and Jansen (2019) used two main sessions of 30 minutes duration on different days and found on average slightly lower reaction times. Meneghetti et al. (2017) used multiple training sessions of approximately 35 minutes duration and found similar reaction time in their second session. Thus, our layout enables the completion of more analyzable tasks in time and achieves the desired increase in power.

Nonlinearity

Although residual plots for models using linear dependency of reaction time on degree show no visual deviations, significant higher order polynomial relationships were observed. Whereas the original paper of Shepard and Metzler (1971) showed perfect linearity in

reaction time, this is not unanimously reproduced, e.g. in Jolicœur, Regehr, Smith, and Smith (1985), and, although often not discussed, inspection of the data in other work rarely shows perfect linearity. Thus, we believe that nonlinearity is not a major concern in discussing if our setup can be used to analyze the same mental processes as the original design. While the linearity allows a simple theoretical framework, namely a fixed mental rotation speed, we have no simple explanation for the nonlinearity. Differing reaction time-accuracy tradeoffs might be one reason, but different strategies analyzing mirrorings instead of rotation might also contribute. In our opinion, the important findings are that reaction time increases and accuracy decreases with increasing angular disparity and deviations from linearity do not show large effects. Until further research is available, we believe, if nonlinearity can be observed in the data, further investigating interactions should be warranted even if they are not significant, e.g. in our data in the case of accuracy differences by side for angular disparity of 180°.

Limitations

One limitation is the inference of mental rotation only from behavioral data. Although it seems plausible, we cannot know if strategies incorporate the usage of both stimuli or if stimuli are actually rotated in the mind from this behavioral data alone and thus cannot infer that the same strategies are used in our design as in the two-figure case. However, the investigation of strategies was not the main goal of our study. Moreover, effects of increasing mental fatigue in time as well as effects of the shape and layout of the cube figures could have influenced our results.

Conclusion and outlook

In conclusion, we have developed a new layout for chronometric mental rotation tests utilizing three figures. While small differences in performance need to be further analyzed,

overall our design is appropriate for analyzing mental rotation performance. The new layout offers an increase of power while also solving possible problems with analysis of accuracy. Our theoretical framework of layout parameters as well as the described stimulus library provide possibilities for future research and implementation.

Using linear mixed models for statistical analysis allowed our work to be one of the first to analyze the progress of participants during the experiment. While we analyzed the absolute improvement in ms and the proportion of correct answers with more time spent on the task, future research could also analyze improvements relative to participants mean performance or with the number of items solved. We also found performance improvements for participants who had experience with mental rotation experiments. There is potential for future work to specify the variable experience more precisely and further explore the relationship to performance.

Further studies could also explore influences on mental rotation performance of different cube figures or different versions of the same figures. This includes the expansion in the three dimensions, the alignment with the origin, and coloring parameters. Design parameters could also be varied in attempt to force different strategies and eye tracking or EEG could be used to analyze the actual use of strategies.

Our setup can also be transferred to mental rotation tasks of the Vandenberg and Kuse (1978) type by adding a mirrored or structurally different base figure to the right of the stimuli, but as in the case of our work, different setups are possible. (Although single item performance in Vandenberg and Kuse tests cannot be tracked and the need is possibly lessened.) Moreover, other pattern recognition tasks could also apply similarly inspired designs.

Declarations

Ethical standards. The experiment was conducted according to ethical declaration of Helsinki. Ethical approval for this study was not required in accordance with the conditions outlined by the German Research Society (DFG) where research that carries no additional risk beyond daily activities does not require Research Ethics Board Approval. We communicated all considerations necessary to assess the question of ethical legitimacy of the study.

Informed Consent. Informed consent was obtained from all individual participants included in the study.

Acknowledgements. We thank Arne Engelhardt (University of Regensburg) for his support with data collection. We thank Philipp Hofmann and Markus Siebertz (University of Regensburg) for their support with testing the stimulus library.

Data availability. The data that support the findings of this study are available at <https://osf.io/dr9mv/> (DOI 10.17605/OSF.IO/DR9MV). The experiment was not preregistered.

Competing interests. The authors declare no competing interests.

Funding. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contributions. All authors contributed to the study conception and design. Material preparation, data collection, and data analysis was performed by Leonardo Jost. The manuscript was written by Leonardo Jost and Petra Jansen. Both authors read and approved the final manuscript.

References

- Baguley, T. (2009). Standardized or simple effect size: What should be reported?. *British journal of psychology*, 100(3), 603-617.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3), 255-278.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
doi:10.18637/jss.v067.i01.
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

- Boone, A. P., & Hegarty, M. (2017). Sex differences in mental rotation tasks: Not just in the mental rotation process!. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(7), 1005.
- Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, 1(1).
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, 39(2), 175-191.
- Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological methods*, 14(1), 43.
- Geary, D. C., Saults, S. J., Liu, F., & Hoard, M. K. (2000). Sex Differences in Spatial Cognition, Computational Fluency, and Arithmetical Reasoning. *Journal of Experimental Child Psychology*, 77(4), 337–353.
<https://doi.org/10.1006/jecp.2000.2594>.
- Hausmann, M. (2014). Arts versus science — Academic background implicitly activates gender stereotypes on cognitive abilities with threat raising men's (but lowering women's) performance. *Intelligence*, 46, 235-245.
<http://dx.doi.org/10.1016/j.intell.2014.07.004>.
- Hedges, L. V. (2007). Effect Sizes in Cluster-Randomized Designs. *Journal of Educational and Behavioral Statistics*, 32(4), 341–370.
<https://doi.org/10.3102/1076998606298043>
- Heil, M., Rösler, F., Link, M., & Bajric, J. (1998). What is improved if a mental rotation task is repeated—the efficiency of memory access, or the speed of a transformation routine?. *Psychological Research*, 61(2), 99-106.
- Hilbert, S., Stadler, M., Lindl, A., Naumann, F., & Bühner, M. (2019). Analyzing longitudinal intervention studies with linear mixed models. *TPM - Testing, Psychometrics, Methodology in Applied Psychology*, 26(1), 101–119.
<https://doi.org/10.4473/TPM26.1.6>
- Jansen-Osmann, P., & Heil, M. (2007). Suitable stimuli to obtain (no) gender differences in the speed of cognitive processes involved in mental rotation. *Brain and Cognition*, 64(3), 217-227.
- Jolicœur, P., Regehr, S., Smith, L. B., & Smith, G. N. (1985). Mental rotation of representations of two-dimensional and three-dimensional objects. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 39(1), 100.

555 Jost, L., Weishäupl, A., & Jansen, P. (2019). Interactions between simultaneous aerobic
556 exercise and mental rotation. Manuscript submitted for publication.

557 Khooshabeh, P., & Hegarty, M. (2010, March). Representations of shape during mental
558 rotation. In *AAAI Spring Symposium: Cognitive Shape Processing*.

559 Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial
560 ability and spatial orientation ability. *Memory & Cognition*, 29(5), 745-756.

561 Krüger, M., & Krist, H. (2009). Imagery and motor processes—When are they connected?
562 The mental rotation of body parts in development. *Journal of Cognition and*
563 *Development*, 10(4), 239-261.

564 Meneghetti, C., Cardillo, R., Mammarella, I. C., Caviola, S., & Borella, E. (2017). The role
565 of practice and strategy in mental rotation training: transfer and maintenance effects.
566 *Psychological research*, 81(2), 415-431.

567 Moè, A. (2016). Does experience with spatial school subjects favour girls' mental rotation
568 performance? *Learning and Individual Differences*, 47, 11-16.
569 <http://dx.doi.org/10.1016/j.lindif.2015.12.007>

570 Moè, A., Jansen, P., & Pietsch, S. (2018). Childhood preference for spatial toys. Gender
571 differences and relationships with mental rotation in STEM and non-STEM students.
572 *Learning and Individual Differences*, 68, 108-115.

573 Nash, J. C., & Varadhan, R. (2011). Unifying optimization algorithms to aid software system
574 users: optimx for R. *Journal of Statistical Software*, 43(9), 1-14.

575 Pek, J., & Flora, D. B. (2018). Reporting effect sizes in original psychological research: A
576 discussion and tutorial. *Psychological Methods*, 23(2), 208.

577 Peters, M., & Battista, C. (2008). Applications of mental rotation figures of the Shepard and
578 Metzler type and description of a mental rotation stimulus library. *Brain and*
579 *cognition*, 66(3), 260-264.

580 Rahe, M., Ruthsatz, V., Jansen, P., & Quaiser-Pohl, C. (2019). Different practice effects for
581 males and females by psychometric and chronometric mental-rotation tests. *Journal*
582 *of Cognitive Psychology*, 31(1), 92-103.

583 Rights, J. D., & Sterba, S. K. (2019). Quantifying explained variance in multilevel models:
584 An integrative framework for defining R-squared measures. *Psychological methods*,
585 24(3), 309.

586 R Core Team (2018). R: A language and environment for statistical computing. R Foundation
587 for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171(3972), 701-703.
- Shepard, S., & Metzler, D. (1988). Mental rotation: effects of dimensionality of objects and type of task. *Journal of experimental psychology: Human perception and performance*, 14(1), 3.
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers*, 31(1), 137-149.
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological bulletin*, 139(2), 352.
- Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and motor skills*, 47(2), 599-604.
- Voyer, D., & Jansen, P. (2016). Sex differences in chronometric mental rotation with human bodies. *Psychological research*, 80(6), 974-984.
- Voyer, D., Jansen, P., & Kaltner, S. (2017). Mental rotation with egocentric and object-based transformations. *The Quarterly Journal of Experimental Psychology*, 70(11), 2319-2330.
- Wohlschläger, A. (2001). Mental object rotation and the planning of hand movements. *Perception & Psychophysics*, 63(4), 709-718.
- Wohlschläger, A., & Wohlschläger, A. (1998). Mental and manual rotation. *Journal of Experimental Psychology: Human perception and performance*, 24(2), 397.
- Xie, F., Zhang, L., Chen, X., & Xin, Z. (2019). Is Spatial Ability Related to Mathematical Ability: a Meta-analysis. *Educational Psychology Review*.
<https://doi.org/10.1007/s10648-019-09496-y>
- Zacks, J. M., & Tversky, B. (2005). Multiple systems for spatial imagery: Transformations of objects and bodies. *Spatial Cognition and Computation*, 5(4), 271-306.

Tables

Table 1. Parameters for generation of cube figures and values used in the experiment.

Parameter group	Parameter	value
Color options	Background color	black
	Border color	black
	Face color	white, grey
Sizing and formatting	Cube Diameter	70px
	Image size	720px*720px
	File format	png
	Centering	none
Model properties	Base orientations	a,b
	Models	Peters and Battista (2008), 1-16
	Base rotation angles (x,y,z)	-30°,0°,15°
	Angle difference	45°
	Rotation axes	y,z

Table 2. Mean (SD) of behavioral data, separated by the side of the congruent base figure (columns) and degree (rows). Reaction time is reported in ms, accuracy in the proportion of correct answers. Mean reaction time is calculated for all correctly answered trials of each participant and then averaged over all participants. Mean accuracy is calculated for all trials of each participant and then averaged over all participants

	Reaction time		Accuracy	
	Left	Right	Left	Right
0°	1824(643)	2074(578)	.97(.04)	.91(.12)
45°	2830(942)	2896(938)	.92(.09)	.92(.09)
90°	3664(1047)	3765(1151)	.87(.11)	.87(.12)
135°	4252(1477)	4343(1467)	.85(.13)	.86(.10)
180°	4804(1467)	4704(1525)	.81(.15)	.86(.11)

Table 3. Statistical analysis of reaction time including axis on the dataset of only rotated stimuli. The values for degree and time (since start of the session) represent estimated changes corresponding to changes of 100° and 30 minutes of testing time. The values for side=right, axis=z, and experience=no are all relative to the other respective condition. The intercept in this model represents the estimated reaction time in ms at time=0, degree=0, side=left, axis=y, experience=yes.

Variable	Estimate	SE	Test statistic	p	95% CI
Intercept	2098.64	201.31			1686.14, 2487.78
Degree*Time	-671.04	98.53	$\chi^2(1)=46.28$	<.001	-863.22, -474.91
Degree*(Side=right)	-113.56	57.00	$\chi^2(1)=3.97$	.046	-223.75, -10.08
Degree*(Axis=Z)	-374.20	56.88	$\chi^2(1)=43.21$	<.001	-481.64, -272.71
Degree*(Time=0)* (Side=left)*(Axis=Y)	2062.09	114.75			1828.19, 2298.92
(Degree=0)*Time	-492.00	159.02			-782.20, -188.61
(Degree=0)*(Side=right)	149.81	63.55			32.34, 269.56
(Degree=0)*(Axis=Z)	387.88	63.46			269.76, 510.26
Experience=no	476.49	203.07	$\chi^2(1)=4.98$	.026	55.34, 862.66
Main Effects:					
Time	-1177.24	123.27	$\chi^2(1)=47.95$	<.001	-1417.51, -908.96
Side=right	34.65	26.41	$\chi^2(1)=1.72$	.190	-21.39, 90.13
Axis=Z	8.14	26.41	$\chi^2(1)=0.10$	.758	-41.39, 61.51
Degree	1457.13	93.63	$\chi^2(1)=78.83$	<.001	1272.79, 1648.52
Nonsignificant Effects:					
Degree*Side*Axis			$\chi^2(1)=0.11$	.746	
Degree*Side*Time			$\chi^2(1)=0.74$	.389	
Side*Axis			$\chi^2(1)=2.55$	.111	
Time*Side			$\chi^2(1)=0.24$	.626	

Table 4. Statistical analysis of reaction time excluding axis on the dataset of all stimuli. The values for degree and time (since start of the session) represent estimated changes corresponding to changes of 100° and 30 minutes of testing time. The values for side=right and experience=no are all relative to the other respective condition. The intercept in this model represents the estimated reaction time in ms at time=0, degree=0, side=left, experience=yes.

Variable	Estimate	SE	Test statistic	p	95% CI
Intercept	2207.46	172.91			1870.43, 2539.50
Degree*Time	-686.97	83.58	$\chi^2(1)=67.33$	<.001	-862.07, -523.55
Degree*(Side=right)	-149.63	48.20	$\chi^2(1)=9.63$	.002	-244.49, -53.33
Degree*(Time=0)* (Side=left)	1994.49	107.73			1787.55, 2192.85
(Degree=0)*Time	-473.26	141.90			-737.29, -195.83
(Degree=0)*(Side=right)	191.26	51.85			91.52, 291.70
Experience=no	323.01	158.03	$\chi^2(1)=3.89$	.049	1.76, 651.41
Main Effects:					
Time	-1125.96	117.44	$\chi^2(1)=48.18$	<.001	-1354.56, -881.59
Side=right	50.00	24.87	$\chi^2(1)=4.04$	.044	6.13, 101.95
Degree	1555.60	94.60	$\chi^2(1)=82.80$	<.001	1359.07, 1755.70
Nonsignificant Effects:					
Degree*Side*Time			$\chi^2(1)=0.32$	.570	
Time*Side			$\chi^2(1)=0.41$	.521	

Table 5. Statistical analysis of accuracy including axis on the dataset of only rotated stimuli using a binomial distribution. The values for degree and time (since start of the session) represent estimated changes corresponding to changes of 100° and 30 minutes of testing time. The values for side=right, axis=z, and experience=no are all relative to the other respective condition. The intercept in this model represents the estimate at time=0, degree=0, axis=y.

Variable	Estimate	SE	Test statistic	p	95% CI
Intercept	3.14	0.25			2.67, 3.65
Degree*(Axis=Z)	0.63	0.11	$\chi^2(1)=31.38$	<.001	0.41, 0.86
Degree*(Axis=Y)	-1.05	0.11			-1.28, -0.84
(Degree=0)*(Axis=Z)	-0.60	0.15			-0.92, -0.32
Time	0.34	0.16	$\chi^2(1)=4.52$	.033	0.02, 0.66
Main Effects:					
Axis=Z	0.12	0.08	$\chi^2(1)=2.06$	.151	-0.04, 0.29
Degree	-0.75	0.09	$\chi^2(1)=45.61$	<.001	-0.93, -0.58
Nonsignificant Effects:					
Side=right	0.10	0.05	$\chi^2(1)=3.55$	.059	-0.00, 0.20
Degree*Side*Axis			$\chi^2(1)=0.57$	.450	
Side*Axis			$\chi^2(1)=0.09$	.761	
Degree*Side*Time			$\chi^2(1)=0.00$	.955	
Degree*Time			$\chi^2(1)=0.93$	.335	
Degree*Side			$\chi^2(1)=3.83$	.050	
Time*Side			$\chi^2(1)=0.88$	.348	
Experience=no	-0.48	0.29	$\chi^2(1)=2.71$	.100	-1.03, 0.07

Table 6. Statistical analysis of accuracy excluding axis on the dataset of all stimuli using a binomial distribution. The values for degree and time (since start of the session) represent estimated changes corresponding to changes of 100° and 30 minutes of testing time. The values for side=right and experience=no are all relative to the other respective condition. The intercept in this model represents the estimate at time=0, degree=0, side=left.

Variable	Estimate	SE	Test statistic	p	95% CI
Intercept	2.99	0.21			2.60, 3.46
Degree*(Side=right)	0.39	0.10	$\chi^2(1)=15.76$	<.001	0.21, 0.61
Degree*(Side=left)	-0.93	0.09			-1.12, -0.74
(Degree=0)*(Side=right)	-0.37	0.11			-0.62, -0.16
Time	0.36	0.16	$\chi^2(1)=4.74$	.030	0.03, 0.70
Main Effects:					
Side=right	0.05	0.05	$\chi^2(1)=1.19$	.276	-0.04, 0.15
Degree	-0.74	0.07	$\chi^2(1)=54.40$	<.001	-0.89, -0.59
Nonsignificant Effects:					
Degree*Side*Time			$\chi^2(1)=0.02$	.879	
Degree*Time			$\chi^2(1)=2.77$	.096	
Time*Side			$\chi^2(1)=0.78$	.378	
Experience=no	-0.40	0.27	$\chi^2(1)=2.04$	.154	-0.92, 0.18

655 Table 7. Statistical analysis of accuracy separated by angle. Intercepts in this model represent
656 the estimate for axis=y, side=left.

Variable	Estimate	SE	Test statistic	p	95% CI
0°					
Intercept	4.31	0.38			3.73, 5.35
Side=right	-1.40	0.31	$\chi^2(1)=22.67$	<.001	-2.18, -0.79
45°					
Intercept	3.06	0.22			2.65, 3.48
Side=right	0.01	0.11	$\chi^2(1)=0.01$	.933	-0.21, 0.24
Axis=Z	-0.15	0.11	$\chi^2(1)=1.72$	.190	-0.39, 0.08
90°					
Intercept	2.33	0.21			1.95, 2.76
Side=right	0.04	0.09	$\chi^2(1)=0.15$	.701	-0.14, 0.21
Axis=Z	-0.19	0.09	$\chi^2(1)=3.98$	.046	-0.38, -0.00
135°					
Intercept	1.90	0.19			1.53, 2.26
Side=right	0.03	0.09	$\chi^2(1)=0.11$	.739	-0.14, 0.21
Axis=Z	0.35	0.09	$\chi^2(1)=15.59$	<.001	0.17, 0.53
180°					
Intercept	1.42	0.20			1.06, 1.85
Side=right	0.43	0.12	$\chi^2(1)=12.48$	<.001	0.21, 0.65
Axis=Z	0.63	0.12	$\chi^2(1)=27.01$	<.001	0.40, 0.88

657

658

Table 8. Estimated coefficients of higher order polynomials for the relationship of reaction time and accuracy with angular disparity at units of 100°. Degree describes the degree of the fitted polynomial and x^1 - x^4 describe the coefficients. Test statistics were computed in comparison to the polynomial of one lower degree.

Dataset	Degree	x^1	x^2	x^3	x^4	Test statistic	p
Reaction time	1	2062.09					
including axis	2	2922.18	-407.6			$\chi^2(1)=35.85$	<.001
	3	4948.50	-2447.06	611.08		$\chi^2(1)=8.66$	.003
Reaction time	1	1994.49					
excluding axis	2	2689.69	-371.33			$\chi^2(1)=68.09$	<.001
	3	2757.85	-464.58	33.41		$\chi^2(1)=0.17$	.682
	4	1863.39	2111.76	-2272.16	637.82	$\chi^2(1)=9.28$	.002
Accuracy	1	-1.05					
including axis	2	-2.57	0.66			$\chi^2(1)=25.29$	<.001
	3	-5.60	3.65	-0.88		$\chi^2(1)=5.19$	.023
Accuracy	1	-0.93					
excluding axis	2	-1.93	0.47			$\chi^2(1)=23.00$	<.001
	3	-1.80	0.30	0.06		$\chi^2(1)=0.10$	.755
	4	0.42	-5.02	4.47	-1.17	$\chi^2(1)=3.80$	.051

a) horizontal layout with vertical shift

b) horizontal layout

c) vertical layout with horizontal shift

666

667 Figure 1. Depictions of different possible layouts of two base figures and one stimulus figure.

668 The stimulus figure is always rotated by 45° around the y-axis for demonstration while the

669 base figures are mirrored to each other.

670

671

672 Figure 2. Depictions of mirror images across the natural mirroring planes. Top left: base
 673 model. Top right: Mirroring across the yz-plane. Bottom left: Mirroring across the xy-plane.
 674 Bottom right: Mirroring across the xz-plane.

675

676 Figure 3. Depictions of rotation by 45° around the natural rotation axes. Top left: Base
 677 model. Top right: Rotation around the x-axis. Bottom left: Rotation around the y-axis.
 678 Bottom right: Rotation around the z-axis.

679

680 Figure 4. Depictions of stimulus figure, that can be rotated into the left base figure by a
 681 rotation of 180° around the y-axis (left) and z-axis(right).