Do the sex differences play such an important role in explaining performance in spatial tasks?

María José Contreras a,⁎, Agustín Martínez-Molina b, José Santacreu c

aFacultad de Psicología, Universidad Nacional de Educación a Distancia (UNED), C/Juan del Rosal, 10, 28040 Madrid, Spain
bFacultad de Ciencias Sociales, Universidad Europea de Madrid (UEM), Villaviciosa de Odón, 28670 Madrid, Spain
cFacultad de Psicología, Universidad Autónoma de Madrid (UAM), Cantoblanco, 28049 Madrid, Spain

1. Introduction

Spatial ability implies the generation, retention, retrieval, and transformation of visual-spatial information. Individuals with a greater spatial ability can form better representations of visual stimuli and they do so faster than individuals with a lower ability; they seem to be more efficient in making the necessary transformations to representations (i.e., rotation) and they automate those processes related to spatial problem-solving quicker (i.e., Lohman, 1988). Since there are numerous different tasks designed to evaluate spatial ability, it is essential to define and isolate the processes with which individuals solve each type of tasks.

Some spatial ability researchers have focused on the issue of gender differences. Previous studies indicate that the largest sex differences are found in spatial skills, where men frequently get a better average performance than women (Neisser, Boodoo, & Bouchard, 1996). The influential meta-analyses carried out by Linn and Petersen (1985) and Voyer, Voyer, and Bryden (1995) on this topic pointed out a greater effect size for tasks that included a Mental Rotation component. It has been stated that the type of university education (science or arts) does not moderate the differences according to sex (Contreras, Colom, Shih, Alava, & Santacreu, 2001); that even though Voyer et al. (1995) suggested that procedural differences are an important source of variance in the magnitude of sex differences, males still have an advantage when these performance factors are partialled out (Contreras, Rubio, Peña, Colom, & Santacreu, 2007). Other studies inform that the solution strategies of a task determine performance in a dynamic spatial task more so than other performance variables but, even in this case, sex differences remain significant (Peña, Contreras, Shih, & Santacreu, 2008; Sacuzzo, Craig, Johnson, & Larson, 1996; Schiff & Oldak, 1990).

Beyond verifying the effect that males outperformed females in overall spatial ability, researchers intend to understand these observed findings. However, researchers have not been able to integrate these results to give a complete psychological explanation for such sex differences. Therefore, our objective is to give a thorough and satisfactory psychological explanation for the sex differences in spatial performance in dynamic tasks. Distinguishing between an instrumental point of view or a theoretical explanation...
on the analyses of sex differences in spatial ability, this study refers to an instrumental view. We hypothesised that when researchers say “sex explains performance on a spatial task”, the sex variable is a construct that includes the cognitive requirements imposed by the task that is being investigated. Therefore, “sex of the individuals” cannot be understood as a causal factor but as an index variable that includes spatial ability. Our research aims at explaining sex differences in spatial task performance in terms of sex differences in lower-level cognitive processes.

1.1. Defining psychological process in solving the SODT task

The Spatial Orientation Dynamic Test (SODT; Colom, Contreras, Botella, & Santacreu, 2002; Colom, Contreras, Shih, & Santacreu, 2003; Contreras et al., 2001, 2007) was developed to assess people’s abilities in dynamic spatial tasks. Previous studies have shown that SODT is a good measure of dynamic spatial ability, tapping into central spatial mental processes (Colom et al., 2002, 2003). A dynamic item represents a setting that changes according to a time function, as does the response. Dynamic tests place emphasis on the psychological processes involved in the task and not just on performance (Grigorenko & Sternberg, 1998). Regarding spatial ability, dynamic tasks are used to test the ability to perceive and extrapolate real motion, to predict trajectories of moving objects, and to estimate arrival times of two or more objects (Colom et al., 2002; Contreras, Colom, Hernández, & Santacreu, 2003; Law, Pellegrino, & Hunt, 1993a; Law et al., 1993b).

The objective of the SODT task is to bring two moving dots as close as possible to a target point. The task performance is measured at the end of the trial by the distance of the dots to the target (DIST). At the beginning of each item of the test, each dot moves at a constant speed and with a predetermined course, moving away from the target point. In order to analyse each process of the SODT task separately and evaluate them independently from the performance task itself, we will describe the processes needed to perform it. The task requires (1) identifying which course button to click; (2) changing the course of the two moving dots in order to direct them to the target which involves (3) managing both the total item time and the time spent on each of the two dots. The three subtasks considered separately correspond, respectively, to the three processes suggested by Peña et al. (2008): Mental Rotation (MR), Spatial Orientation (SO) and Time Management (TM). The indexes that measure each process are detailed in Section 2.3. Fig. 1 shows a screen shot that represents an item of SODT.

Peña et al. (2008) suggest that MR, SO and TM are involved in SODT performance. Time Management is crucial in numerous everyday activities and essential in certain jobs, such as air traffic controlling. It should, in any case, be taken into account when dealing with dynamic Spatial Orientation tasks such as SODT, as it requires manoeuvring two moving dots for each item in a limited amount of time.

We designed a new version of SODT (2.0 v) which allows us to measure each of these subtasks separately as well as the overall performance of the task. The objectives of this study are (1) to examine whether subtasks indexes are independent of each other and (2) to determine the predictive power of each one of them on task performance. We expect that performance in the SODT task (DIST) will show significant differences between males and females, as previous studies have shown. But we also expect that the three variables obtained in the new version of SODT (MR, SO, and TM) together will explain most of the task performance variance. Consequently, sex will then lose predictive power in favour of other variables: we hypothesised that sex differences in spatial task completion occur because there are sex differences in the processes that measure the task. The predictive power of these processes in task execution is as expected since (a) MR and SO are clearly implicated in spatial task performance (i.e. Lohman, 1988; Voyer et al., 1995) and (b) TM seems to be a critical variable in dynamic tasks because the item changes according to a time function (Grigorenko & Sternberg, 1998) and the response time for each trial is limited.

2. Method

2.1. Participants

In this study, the participants were 341 candidates of a selection process applying for a (training) scholarship for a highly technical job. All the participants had a higher-level education (diploma or degree), as this was a requirement for the scholarship. Out of the total amount of participants, 225 were males (66%) and 116 were females (34%). The mean age of the females was 28.9 years (SD 3.3), whereas the mean age of the males was 29.9 years (SD 3.8). Although we must acknowledge that the sample is not representative of the population, the important point is that this is true for both males and females.

2.2. Measurements and procedure

Participants completed the SODT 2.0 test in a single 3 h (approx.) session, where a battery of 14 psychological tests was applied. Six of them were measures of diverse cognitive abilities and eight were personality tests. These tests were all computerised within an evaluation system in a web environment. The computerised system supplied the different tests on the evaluation computers and the data generated throughout the process were collected in a server controlled by the evaluating team.

The SODT 2.0 test is a version of a number of SODT tasks (the task has been widely discussed by Contreras et al. (2007) which allows one to obtain data on MR, on dynamic SO and on TM for each task. The performance in this spatial task is the distance (DIST) at which the differently coloured moving dots end up from the target. Each trial presents two moving dots of different colours on the computer screen that have to be directed towards a target point situated approximately in the centre of the screen. In each trial, the moving dots start in different positions, equally distant from the target point and at the same speed, with an initial course predetermined at $0^\circ$, $90^\circ$, $180^\circ$ or $270^\circ$. A control panel situated on the top of the screen manipulates each moving dot (see Fig. 1). The control panels are in the same colour as the moving dot they control (blue or red). The course of the dots can be modified by...
pressing the adjacent black arrows. Each press modifies the course of the moving dot by 10°.

The task consists in bringing the moving dots to the destination, adjusting the trajectory of each moving dot so as to direct it towards the target point as quickly as possible. This is accomplished by repeatedly pressing on the buttons on the control panel until both dots get as close as possible to the target.

The test is comprised of 13 trials. The first four trials are considered as training and the last nine are the evaluation trials, from which the indexes explained below are calculated. Each evaluation trial lasts 25 s.

In previous studies with the SODT test, internal consistency indexes ranged around .9 for the SO variable (e.g. Contreras et al., 2007); from .9 to .95 for the DIST variable (Colom et al., 2003; Peña et al., 2008) and of .92 for the MR variable (Martínez-Molina, 2008).

### 2.3. Variables

**Task performance (DIST):** Performance on the spatial task corresponds with the objective of such a task: to get as close as possible to the target point. It is measured by the mean distance, expressed in pixels, from each one of the dots to the target at the end of each trial. It is considered a variable dependent on the other three that the test provides.

**Mental Rotation (MR):** This is the first subtask at the beginning of each item (display). MR can be measured by the individual’s first click over the controls. The participant must choose which direction button click to change the course of each of the dots. In SODT test, MR is to reduce the Angular Discrepancy (AD) between the course of the moving dot and the course necessary to reach the target, but considering their relationship with the orientation of the arrow button on the panel control. MR is expressed as a ratio: Sum of the AD performed/Sum of the AD configured. The possible MR score for each item is 0, 1, 2. The MR score depends on the participant’s first click on the buttons that correct the course of each of the two moving dots. If this first click moves both dots away from the target, the score is 0. If this first click only moves one of the dots closer to the target, then the score is 1 and if the first click moves both dots towards the target, then the score is 2.

**Spatial Orientation (SO):** This variable is measured at the end of the item by the deviation of each moving dot. The deviation of each dot is the difference between the course of the dot and the correct course towards the target. Participants must modify the trajectory of both moving dots by clicking on the corresponding arrows on the control panel for each dot. The task is carried out by modifying each dot’s trajectory and observing its course throughout the trial. SO is the mean deviation for each moving dot in the complete set of trials. If deviation is equal 0°, the moving dot is on the correct direct course to the target.

**Time Management (TM):** It is important that the participants manage the time they have and that they quickly correct the course of both dots. Otherwise while one of the dots moves towards the target, the other, maintaining its original course, will move away from the target (see Fig. 2). Keep in mind that the task performance is measured by the distance from each one of the dots to the target at the end of each trial. TM is expressed in seconds and it corresponds to the moment in which the individual, after modifying the course of the first dot, starts to modify the second dot.

### 2.4. Analyses

Firstly, the descriptive statistics were calculated according to sexes. The objective was to corroborate that, within the sample, there are significant mean differences between the groups of males and females using DIST as a measure of spatial performance. Afterwards, the correlation between the variables from the three processes (MR, SO, TM) and the task performance (DIST) was studied. Finally, a competitive, multiple linear regression model was carried out in which, apart from the variables MR, SO, and TM, the variable sex (dummy-coded variable, 0 = male, 1 = female) was introduced, considering performance in the SODT 2.0 test as the dependent variable. The objective was to analyse the relevance of the sex variable when it competes with other variables from the SODT 2.0 task to explain performance on a spatial task.

### 3. Results

#### 3.1. Descriptive statistics and correlational analyses

Table 1 shows the descriptive statistics for the male and female samples separately. It is observed that the group of males, take less time to move onto the second moving dot (TM variable). They, on average, also rotate correctly at the start of the task, and they orientate the dots better. In addition, they are more efficient, as they leave the moving dots closer to the target. The comparison between means is significant for all variables and the effect size is high¹ for the variable DIST and medium for TM, MR and SO.

Table 2 shows the correlations obtained between the different variables analysed. The correlation between the dependent variable and each subprocess is significantly high (please see values in the first three rows in Table 2). The correlation between DIST and TM presents a high value (.65) in the expected direction: the sooner the first dot is orientated and the quicker the change to the second dot, the more efficient the performance in the task will be. DIST has a significant relationship with MR (−.54), indicating that the better the rotation at the beginning of the task, the less scoring for distance is obtained in the trial.

As was expected, the relationship between MR and TM is low (−.16). Both variables are considered independent. MR has to do

---

¹ Cohen (1988) suggested that values of $d = .2$ are considered small, $d = .5$ medium and $d = .8$ or higher are considered large.
with thinking in “how to” orientate the dots and is measured with the first press, while the time that is dedicated afterwards to efficiently direct the dots would not have anything to do with such MR exercise.

The correlation between TM and SO is notably lower (.26) than that for DIST although it is also positive and significant, as was expected. This less intense relationship is rationally understandable, as the SO index is not as related to the time in which the task is solved, unlike DIST, which is directly related to the time it takes to orientate the moving dots correctly. An individual can make the dot turn several times and still end up directing it correctly, which would imply a low deviation index (SO), but with a greater distance to the target (DIST) than if the dot had been directed correctly sooner, avoiding the dot moving further away.

When correlation values for males and females are compared separately, a similar pattern is observed for both groups, which is directly sooner, avoiding the dot moving further away. An individual can make the dot turn several times and still end up directing it correctly, unlike DIST, which is directly related to the time it takes to orientate the moving dots correctly. An individual can make the dot turn several times and still end up directing it correctly, which would imply a low deviation index (SO), but with a greater distance to the target (DIST) than if the dot had been directed correctly sooner, avoiding the dot moving further away.

When correlation values for males and females are compared separately, a similar pattern is observed for both groups, which is directly sooner, avoiding the dot moving further away. An individual can make the dot turn several times and still end up directing it correctly, unlike DIST, which is directly related to the time it takes to orientate the moving dots correctly. An individual can make the dot turn several times and still end up directing it correctly, which would imply a low deviation index (SO), but with a greater distance to the target (DIST) than if the dot had been directed correctly sooner, avoiding the dot moving further away.

When correlation values for males and females are compared separately, a similar pattern is observed for both groups, which is directly sooner, avoiding the dot moving further away. An individual can make the dot turn several times and still end up directing it correctly, unlike DIST, which is directly related to the time it takes to orientate the moving dots correctly. An individual can make the dot turn several times and still end up directing it correctly, which would imply a low deviation index (SO), but with a greater distance to the target (DIST) than if the dot had been directed correctly sooner, avoiding the dot moving further away.

When correlation values for males and females are compared separately, a similar pattern is observed for both groups, which is directly sooner, avoiding the dot moving further away. An individual can make the dot turn several times and still end up directing it correctly, unlike DIST, which is directly related to the time it takes to orientate the moving dots correctly. An individual can make the dot turn several times and still end up directing it correctly, which would imply a low deviation index (SO), but with a greater distance to the target (DIST) than if the dot had been directed correctly sooner, avoiding the dot moving further away.

When correlation values for males and females are compared separately, a similar pattern is observed for both groups, which is directly sooner, avoiding the dot moving further away. An individual can make the dot turn several times and still end up directing it correctly, unlike DIST, which is directly related to the time it takes to orientate the moving dots correctly. An individual can make the dot turn several times and still end up directing it correctly, which would imply a low deviation index (SO), but with a greater distance to the target (DIST) than if the dot had been directed correctly sooner, avoiding the dot moving further away.

When correlation values for males and females are compared separately, a similar pattern is observed for both groups, which is directly sooner, avoiding the dot moving further away. An individual can make the dot turn several times and still end up directing it correctly, unlike DIST, which is directly related to the time it takes to orientate the moving dots correctly. An individual can make the dot turn several times and still end up directing it correctly, which would imply a low deviation index (SO), but with a greater distance to the target (DIST) than if the dot had been directed correctly sooner, avoiding the dot moving further away.

When correlation values for males and females are compared separately, a similar pattern is observed for both groups, which is directly sooner, avoiding the dot moving further away. An individual can make the dot turn several times and still end up directing it correctly, unlike DIST, which is directly related to the time it takes to orientate the moving dots correctly. An individual can make the dot turn several times and still end up directing it correctly, which would imply a low deviation index (SO), but with a greater distance to the target (DIST) than if the dot had been directed correctly sooner, avoiding the dot moving further away.

When correlation values for males and females are compared separately, a similar pattern is observed for both groups, which is directly sooner, avoiding the dot moving further away. An individual can make the dot turn several times and still end up directing it correctly, unlike DIST, which is directly related to the time it takes to orientate the moving dots correctly. An individual can make the dot turn several times and still end up directing it correctly, which would imply a low deviation index (SO), but with a greater distance to the target (DIST) than if the dot had been directed correctly sooner, avoiding the dot moving further away.

When correlation values for males and females are compared separately, a similar pattern is observed for both groups, which is directly sooner, avoiding the dot moving further away. An individual can make the dot turn several times and still end up directing it correctly, unlike DIST, which is directly related to the time it takes to orientate the moving dots correctly. An individual can make the dot turn several times and still end up directing it correctly, which would imply a low deviation index (SO), but with a greater distance to the target (DIST) than if the dot had been directed correctly sooner, avoiding the dot moving further away.

When correlation values for males and females are compared separately, a similar pattern is observed for both groups, which is directly sooner, avoiding the dot moving further away. An individual can make the dot turn several times and still end up directing it correctly, unlike DIST, which is directly related to the time it takes to orientate the moving dots correctly. An individual can make the dot turn several times and still end up directing it correctly, which would imply a low deviation index (SO), but with a greater distance to the target (DIST) than if the dot had been directed correctly sooner, avoiding the dot moving further away. The first model shows that SO (deviation) has a great predictive power over the task performance (DIST). The second model includes the TM variable also, which increases the predictive power by $\Delta R^2 = .220$. The third model adds MR variable (AD performed/AD config.) to the previous model, which also increases the predictive power of the model to $R^2 = .073$. Finally, in the fourth model, Sex contributes to the explanation with a non-significant value ($p = .635, > .05$) without a predictive increment ($\Delta R^2 = .000$).

The model fulfills the assumptions that guarantee the independence of the variables. The value of the Durbin–Watson was 2.074, indicating that the residuals are completely independent.

It also satisfies the assumption of homoscedasticity. At least, the distribution of the residuals does not show a linear relationship. There is no established linear relationship between the standardised predicted values and the standardised residuals ($r = .000, p < .001$).

Finally, indicators show no multicollinearity problems (Tolerance .877, .913, and .952; Variance Inflation Factors 1.140, 1.095, and 1.050).

4. Discussion

This study has two main objectives. The first one was to analyse the relationship between the variables that measure the processes of the SODT 2.0 task and their relationship with performance in such tasks. Together, TM and SO explain a great amount of variance in distance ($R^2 = .798$), with an increase of .220 with respect to the model that only included the SO variable. When a model that adds TM to variables related to spatial processes (MR and SO) is proposed, the regression model explains the final performance in the dynamic task with a notable explained variance of $R^2 = .871$. These three variables (MR, SO, TM) can almost completely explain the overall task performance (DIST).

The second objective of this study was proposed in relation to an old problem in spatial tasks in general, and in the dynamic task SODT in particular. It involves the search for, through numerous studies, the variables that could explain sex differences, systematically in favour of males (as a group), when solving this task. Contreras et al. (2001) proved that the differences in mean performance between groups were not related to the type of university studies carried out by the participant. Contreras et al.’s (2007) study on performance could not explain the differences shown by males and females when solving the SODT task. However, a clue to solving this problem was found when the differences amongst variables were discovered: males took longer to press for the first time, which led to the hypothesis that it could be related to the time taken for the MR process. In the present study, it has been demonstrated that males have a greater ratio of MR, which adds plausibility to such a hypothesis. Moreover, other differences in performance factors in Contreras et al. (2007) also found that males had a lower overall response frequency, a greater amount of correct responses and that the time until the last press was less

Please cite this article in press as: Contreras, M. J., et al. Do the sex differences play such an important role in explaining performance in spatial tasks? Personality and Individual Differences (2012), doi:10.1016/j.paid.2011.12.010
(they considered that the dots were correctly orientated before the group of females). Even though these variables alone did not reduce the sex differences, they indicated that males had a more efficient behaviour throughout the task (less clicks, more correct clicks and they do it in less time). This means that the sex differences is not the variable that helps to solve the task correctly but rather using a certain type of process that is efficient in solving spatial tasks (Casey, Nuttall, & Pezaris, 1999; Peña et al., 2008).

To check this hypothesis, different regression models were analysed, the last one of these included the sex variable, in addition to MR, SO and TM. The results showed that the explicative power of the sex variable in the prediction of efficient performance was practically null. However, the differences in the means and the effect size in the male and female samples were very significant. The data presented in this study add the “sex” variable last to the regression model but, additionally, another analysis was performed, starting with a regression model that included sex as its first variable. Results obtained were similar, the predictive power of sex in this model was $R^2 = .15$. When the model included other variables like MR and TM, the predictive value increased considerably ($R^2 = .42$).

The results of this study are clear: the processes carried out during each trial determine participants’ efficiency in solving a dynamic spatial task (SODT 2.0). Firstly, modifying quickly the course of each one of the two moving dots, without allowing them to get too far away, explains some of the scoring in distance; secondly, aiming correctly towards the target and, finally, reducing the Angular Discrepancy at the start of each trial. The fact that there are more males than females that perform these processes correctly is related to some classical findings from the literature on spatial tasks. Voyer et al. (1995), in their meta-analysis study, found that the magnitude of the advantage that males have on females depended on the type of spatial task. The greatest difference was found with the rotation tasks, followed by spatial perception and, finally, by visualization.

On the other hand, Contreras et al. (2001) demonstrated that spatial tests require the same cognitive ability in males and females. While it is true that males outperform females in spatial tests, this is not the same as saying that spatial tests do not measure the same spatial ability in both sexes. Contreras et al. (2001) computed the congruence coefficients for several factors representing general spatial ability, dynamic spatial performance, static spatial performance, and reasoning. The obtained values were +.984, +.959, +.82, and +.84, respectively. Those results show that SODT does measure the same underlying ability, irrespective of sex. Thus, the high number of males in the present study can hardly explain the observed results. In a similar way, Contreras et al. (2007) showed that the patterns of performance for the group of males and the group of females were similar.

Finally, the studies performed with the SODT task on the effect of sex converge in the same result: only the absolute levels of performance differ between sexes.

Acknowledgements

This work has been partially funded by the Ministerio de Ciencia y Tecnología (Science & Technology State Department) (SEJ2007-64448/PSIC) and by the Scientific-Technical Collaboration Agreement AENA-UAM/785005. The authors thank both Reviewers for their conscientious and stimulating comments, which have greatly enriched this work.

References