

THE TRANSITIVITY MISCONCEPTION OF PEARSON'S CORRELATION COEFFICIENT

ANA ELISA CASTRO SOTOS
Katholieke Universiteit Leuven
anaelisa1980@gmail.com

STIJN VANHOOF
Katholieke Universiteit Leuven
stijn.vanhoof@ped.kuleuven.be

WIM VAN DEN NOORTGATE
Katholieke Universiteit Leuven
wim.vandennoortgate@ped.kuleuven.be

PATRICK ONGHENA
Katholieke Universiteit Leuven
patrick.onghena@ped.kuleuven.be

ABSTRACT

Despite the relevance of correlational studies for most research domains, many students, teachers, and researchers alike hold misconceptions concerning the Pearson product-moment correlation coefficient. One of these, the transitivity misconception, has not yet been documented in a systematic way. This paper summarizes the first empirical study, using 279 university students, and examines the relationship between student-based and task-based factors and the appearance of this misconception. In particular, two task-based factors seemed to have a significant effect on its appearance. In addition, the respondents' level of confidence in their answer to the transitivity item was significantly lower than for most other times.

Keywords: *Statistics education research; University students; Confidence*

1. INTRODUCTION

Since the early 1970s, research has detected and described typical inaccuracies in ordinary reasoning involving chance and randomness, documenting classic errors concerning probability and statistics (e.g., Kahneman, Slovic, & Tversky, 1982; Konold, 1989, 1991; Shaughnessy, 2003). With regard to misconceptions of correlation, since the early studies of Inhelder and Piaget (1955), most psychological research has suggested that, just like probabilistic reasoning, human associative reasoning is generally very poor (Nisbett & Ross, 1980). Most misconceptions studied in this literature are related to the way data are presented (verbally, in a table, through plots, or by means of correlation coefficients; Batanero, Estepa, Godino, & Green, 1996), the question that must be answered, and/or the type of variables between which correlation is judged (Beyth-Marom, 1982). Based on their empirical studies, Batanero et al. (1996) and Estepa and Sánchez Cobo (2001, 2003) built a classification for students' most common misconceptions about correlation:

- *Causal*
Students confound association with causality and believe that correlation can always be attributed to a causal relationship.
- *Deterministic*
Students expect correlated variables to be linked by a mathematical function such as a proportion or a power function.
- *Local*
Students use only part of the data provided in a contingency table, neglecting relevant information that might appear in other cells.
- *Unidirectional*
Students perceive the association only when the sign is positive and they consider a negative correlation as independence.

In our experience with students of statistics we have become aware of another misconception concerning correlation that has not been empirically documented so far. It concerns the non-transitivity property of the correlation. This characteristic of the correlation has been studied from a mathematical and a statistical point of view, but not yet from the educational perspective.

Unlike many common mathematical binary relations (e.g., equality, being greater than, divisibility), correlation is not transitive. It is not difficult to find examples of non-transitive relations in real life. For instance, “being friends with” is not a transitive relationship (e.g., Mary is friends with Jane and Jane is friends with Paula, but Mary might not be friends with Paula!). However, as opposed to this type of example outside of daily experience, mathematical binary relations taught along school grades are mostly transitive, which makes correlation an exception and its non-transitivity property prone to be neglected by students.

As already pointed out by McNemar (1969), the non-transitivity property of the correlation implies that, given three quantitative random variables X , Y , and Z , such that X and Y as well as Y and Z are positively correlated (in terms of Pearson’s correlation coefficients, $\rho_{XY} > 0$ and $\rho_{YZ} > 0$), X and Z are not necessarily positively correlated (pp. 154-187). In fact, X and Z might be uncorrelated ($\rho_{XZ} = 0$) or even negatively correlated ($\rho_{XZ} < 0$).

More recently, Langford, Schwertman, and Owens (2001) showed that the correlation between X and Z certainly falls into the following interval:

$$\rho_{XY}\rho_{YZ} - \sqrt{(1-\rho_{XY}^2)(1-\rho_{YZ}^2)} \leq \rho_{XZ} \leq \rho_{XY}\rho_{YZ} + \sqrt{(1-\rho_{XY}^2)(1-\rho_{YZ}^2)} \quad (1)$$

After some algebra, it can be deduced from Equation 1 that transitivity (the certainty that $\rho_{XZ} > 0$) for positive correlations ($\rho_{XY} > 0$ and $\rho_{YZ} > 0$) only holds under very restrictive conditions, more specifically, when they satisfy the following property (Condition 2):

$$\rho_{XY}^2 + \rho_{YZ}^2 > 1 \quad (2)$$

When looking for positive values of these two correlations that fulfill the condition, it is easy to see that several combinations of dissimilar values, such as 0.96 and 0.25, would suffice. However, when searching for the lowest possible values for identical correlations satisfying this condition, values at least as high as $\sqrt{0.50} = 0.707$ are necessary.

An alternative way of looking at the non-transitivity property of the correlation is by squaring the correlation coefficients and making a graphical representation of the situation. The square of a Pearson's coefficient is the proportion of variance in one variable explained by the other (or the proportion of shared variance between the two). By representing these proportions of variance via overlapping squares (see Figure 1 for $\rho_{XY} = \rho_{YZ} = 0.5$ and therefore $\rho_{XY}^2 = \rho_{YZ}^2 = 0.25$), it is easier to see that even if both X and Z are positively correlated with a variable Y , X and Z can be uncorrelated, as shown in Figure 1 by the empty intersection of X and Z . This approach also clarifies Condition 2: If the sum of the proportion of the variance of Y that is shared with X and the proportion of the variance that is shared with Z is larger than 1, at least part of the variance of Y must be shared with both X and Z . In the graphical representation, this would result in a non-empty intersection of X and Z .

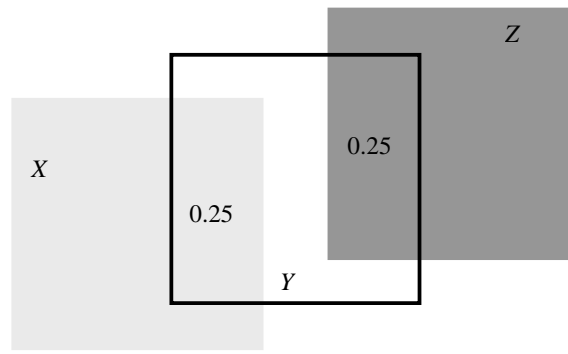


Figure 1. Non-Transitivity for the Proportion of Variance when $\rho_{XY} = \rho_{YZ} = 0.5$

In many science domains such as education, correlations of a magnitude that satisfy Condition 2 are very rare. The widely accepted standards established by Cohen (1988), for example, consider correlation coefficients of 0.1, 0.3, and 0.5 as small, medium, and large respectively, and not one of those values, even up to 0.7, satisfies the necessary Condition 2 above. The mathematical boundary for transitivity is therefore much higher than correlations that can be expected in practice for numerous areas of research.

Real-Data Example of the Non-Transitivity For any positive values not satisfying Condition 2, the correlation between X and Z is allowed to fall into an interval, that, although primarily positive, includes zero as well as negative values. For instance, if $\rho_{XY} = \rho_{YZ} = 0.7$ (and therefore $\rho_{XY}^2 + \rho_{YZ}^2 = 0.98 < 1$), ρ_{XZ} falls between -0.02 and 1 . A real case scenario where the correlation between X and Z effectively takes one of such values was reported by Moore (2006), who found three published variables related to film ratings showing non-transitivity. Namely, *rating* and *length* of the movie were positively correlated ($\rho_{XY} = 0.32$, $\rho_{XY}^2 = 0.10$), as well as *length* of the movie and *year* of release ($\rho_{YZ} = 0.51$, $\rho_{YZ}^2 = 0.26$). However, *rating* and *year* of release were found to be negatively correlated ($\rho_{XZ} = -0.15$).

For the present research, we defined the *transitivity misconception* of the correlation as the belief that given two positive correlations between variables X and Y , and Y and Z that do not satisfy Condition 2, the correlation between X and Z will certainly be positive.

This belief contrasts with the non-transitivity property of the correlation characterizing that, for example, for two initial correlations smaller than 0.7 (which are most common in real settings), the third correlation might be zero or even take negative values.

2. RESEARCH AIMS AND QUESTIONS

As mentioned above, the transitivity misconception of the correlation has not been empirically documented so far from an educational viewpoint. The study presented here intended to fill that void and focused on university students following introductory statistics courses in order to explore their knowledge, confidence, and possible factors affecting them when confronted with a transitivity problem. There were mainly three aims motivating this study:

1. Surveying to what extent students of introductory statistics commit the transitivity misconception.
2. Identifying the extent to which student- and/or task-based factors affect this performance.
3. Exploring the degree of confidence that students have in their answers to the transitivity problem, and its relation to the presence of the misconception and/or the factors studied in the second aim.

Each of these three interests led us to a few specific research questions. Concerning the first and second aims, we acknowledge that considering correlation to be transitive (just like previously learned mathematical relations) gives the correct conclusions in extreme cases, when the correlations ρ_{XY} and ρ_{YZ} are large enough such that Condition 2 holds (one of them extremely high or both higher than $\sqrt{0.50}$). However, in more realistic cases, where these correlations are not that high, we suspect that a transitivity misconception will arise as a consequence of extrapolating the same idea for any positive ρ_{XY} and ρ_{YZ} . Hence, we constructed the following research questions to address our first two goals:

- 1a. How common is the transitivity misconception? In other words, how many of our participants choose that the correlation between X and Z will be positive in cases where this is not necessarily true?
- 1b. How strong is this belief? More specifically, from all participants showing evidence of the misconception, how many merely select that the correlation between X and Z will be greater than 0, and how many that it will be even higher (higher than 0.5)?
- 1c. Do students also have a correct intuition? Namely, regardless of the presence of the misconception and its strength, how many of our participants choose that the correlation between X and Z can be 0 or smaller?
- 2a. Are any students' characteristics (age, field of study, year, gender, and statistical background) related to the appearance of the transitivity misconception?
- 2b. Do the three following task-based factors have an effect on the appearance of the transitivity misconception?
 - Type of representation
 - Size of correlations between X and Y , and Y and Z
 - Contextualization of the problem
- 2c. Do those factors have an effect on the strength of the misconception?
- 2d. Do those factors have an effect on the intuition that the correlation between X and Z can be 0 or negative?

With regard to the third aim, we focused our methodology and data analyses on answering the following research questions:

- 3a. What is the average confidence of participants in their answer to the transitivity item as compared to other items?
- 3b. Does the degree of confidence in the answer differ significantly between students falling into the misconception and those showing no evidence of it?
- 3c. Are any students' characteristics (age, field of study, year, gender, and statistical background) related to the degree of confidence in the answer to the transitivity item?
- 3d. Do the three task-based factors have an effect on the degree of confidence in the answer to the transitivity item?

3. METHODOLOGY

In the context of a broader study about statistical misconceptions, a between-subjects design was applied for a five-item multiple-choice questionnaire distributed in January 2007 to 279 university students (188 females, 87 males, 4 unknown) from different disciplines (Mathematics, Statistics, Psychology, Medicine, and Business) who were taking introductory statistics courses at the Complutense University (Madrid, Spain). Most participants were in their first year (71%) and had never taken a university statistics course before (65%), forming a quite homogeneous group. The questionnaire was voluntarily completed and handed in during class time at the end of the semester in the presence of the teacher, as an anonymous exercise.

In order to find complementary information about the motivation of students and the reasoning behind their answers and confidence level, and to help us frame the quantitative results obtained from the responses to the questionnaire, we invited all participants to take part in a semi-structured interview with one of the authors. Nine students volunteered for it. More detailed information about the questionnaire and the interview can be found in the following two subsections.

3.1. THE QUESTIONNAIRE

All students participating in this research were taking an introductory statistics course that covered the topic of correlation (but not explicitly the non-transitivity property) and therefore were potentially qualified for answering a specific item about transitivity. Two types of questionnaire were constructed including such an item and four extra questions that depended on the rest of the content of the course and the research goals for the broader study. First, for students enrolled in a course that did not cover the topic of hypotheses tests, the questionnaire focused only on features and typical misconceptions around the idea of correlation, as follows:

Item 1.1: About the commutative property of the correlation

Item 1.2: The transitivity item

Item 1.3: About the causality misconception

Item 1.4: About the relation between Pearson's coefficient and the proportion of variance

Item 1.5: About the causality misconception

Second, for students taking a course including the topic of hypotheses tests, the questionnaire comprised the next items:

Item 2.1: The transitivity item (same as 1.2)

Item 2.2: About the causality misconception (same as 1.5)

- Item 2.3: About the definition of hypotheses tests
 Item 2.4: About the interpretation of a p-value
 Item 2.5: About the interpretation of the significance level

The transitivity item (1.2 and 2.1) addressed the research questions that motivated our study and therefore, this paper focuses only on the results concerning this question. The other items constituting the two versions of our instrument can be found in the Appendix.

The Transitivity Item The transitivity item presented participants with the correlations between (or proportion of variance explained by) X and Y , and Y and Z , and asked them to select conclusions about the correlation (or proportion of variance) between X and Z out of several given options. For both versions of the questionnaire, 12 different variations of this item were constructed (and randomly assigned) according to three factors. The factors and their levels were based on the theoretical considerations described in the Introduction and are shown in Table 1.

Table 1. Factors and levels for transitivity item

Factor	Levels
Type of Representation	1. Pearson's correlation coefficient 2. Proportion of variance
Size of Correlations ρ_{XY} and ρ_{YZ} (ρ_{XY}^2 and ρ_{YZ}^2 for proportion of variance version)	1. 0.3 (0.09 for proportion of variance version) 2. 0.5 (0.25 for proportion of variance version) 3. 0.7 (0.49 for proportion of variance version)
Contextualization	1. Contextualized problem 2. Theoretical wording
TOTAL: 3 Factors	TOTAL: $2 \times 3 \times 2 = 12$ Levels

The interest in the two types of representation comes from the idea illustrated above that a mental representation in Venn-like diagrams when working with proportions of variances could help the understanding of the problem. Concerning the values for the correlation, representative values were selected for ranges intuitively low (0.3 is medium for Cohen, 1988), medium (0.5 is large for Cohen, 1988), and large (0.7) respectively. Finally, with regard to the contextualized problem, a cover-story was used that referred to grades in Mathematics, English, and French (variables X , Y , and Z respectively in the theoretical version). The different options that students could select as answers to the item are listed in Table 2.

Table 2. Response options in the transitivity item

Option	Pearson's coefficient version	Proportion of variance version
A	It will be positive	It can be 0
B	It can be 0	It will be higher than 0.25
C	It will be higher than 0.5	It will be higher than 0.5
D	It can be negative	It will be between 0 and 1
E	It will be between -1 and 1	-

In summary, the contextualized version of the item looked as follows. Note that all possible options are indicated (separated by forward slashes "/") but only one was (randomly) assigned to each student. The information between brackets refers to the options for the proportion of variance version of the item.

3.2. THE INTERVIEW

With the completed questionnaire at hand and in a one-to-one context, one of the authors conducted a semi-structured interview with those students who volunteered to participate. The interviews took about 15 minutes on average, were audio-taped and later on transcribed and translated. Transcriptions are available from the authors upon request. In addition, the interviewer took field notes on what the interviewees said that helped in following the dialogues in the audio tapes.

In the initial phase (first five minutes approximately), the interviewer gathered general information about three issues that helped us putting the responses of the student in context. First, the student was asked to talk openly about his or her perception of the questionnaire's difficulty. Second, the student was asked to point out which concepts in the questionnaire were unfamiliar for him or her and in what way. And third, the interviewer asked the student about his or her reasons for taking the present statistics course.

The rest of the interview was devoted to the exploration of the reasoning behind the answers to each of the items, and the justification for the chosen confidence level, especially when it differed appreciably across items. This way, we could get more insight into the answers to our research questions than the statistical analysis of the questionnaires would suggest. In other words, the interviews provided us with extra information about the reasons why this misconception was (or was not) common among students (first research aim), to what extent participants were aware of significant factors affecting their performance (second research aim), and why specific levels of confidence were (or were not) related to correct (or incorrect) answers (third research aim).

This structure for the interviews was taken merely as a guideline and extra space was given for personal explanations, discussions, and questions for each participant according to her/his characteristics, responses, and the natural evolution of the conversation.

4. RESULTS FROM THE QUESTIONNAIRE

The following three subsections present the results extracted from the participants' responses to the questionnaire concerning the research questions proposed in Section 2.

4.1. PRESENCE OF THE TRANSITIVITY MISCONCEPTION

In order to address our first research question, *How common is the transitivity misconception among our students of university introductory courses of statistics?* (1a), we classified participants' responses to the transitivity item by means of a simple coding (Coding 1) based on the answer options to the transitivity item (Table 2). If a student selected, at least, options a or c (b or c for the proportion of variance version) the answer was coded "1" (the misconception is present). On the other hand, if the student did not select options a nor c (neither b nor c in the proportion of variance version) the answer was coded "0" (no evidence of the misconception).

In addition, we considered as well a more refined classification of students' answers (Coding 2) that distinguished two subcategories for answers not showing evidence of the misconception. This coding assumed the same combination of responses as Coding 1 for the presence of the misconception (a or c, b or c for the proportion of variance version) and two subcategories for answers not showing evidence of the misconception. The first subcategory corresponded to completely correct answers (coded as "0") when the student selected exactly b, d, and e, and did not select a nor c (not b, not c, a, and d for the proportion of variance version). The second category included other errors (coded as "2")

that were used when any other combination of responses not fitting in the previous two categories was provided.

According to these coding systems, almost half (49%) of the participants in this study (137 out of $n = 279$, percentage of missing data: 3%) showed evidence of the transitivity misconception by selecting options a or c (b or c in the proportion of variance version) for the transitivity item. In other words, almost half of the students filling in the questionnaire selected that the correlation between two variables (X and Z) will be positive, or even higher than 0.5, when given positive correlations between these two variables and a third one (Y). Furthermore, using Coding 2, we found out that only 13% of the participants (37 out of $n = 279$) provided a completely correct answer.

Regarding our second research question, *How strong is the misconception?* (1b), a further look at the possible answers to the transitivity item revealed that participants showing evidence of transitivity held a *soft* misconception more often than a *strong* misconception. On the one hand, 63% of them (86 out of $n = 137$) chose option a but not c (b but not c for the proportion of variance version), which is equivalent to affirming that the correlation between X and Z will be higher than 0 (0.25 in the proportion of variance version) but not higher than 0.5. On the other hand, the rest of them took the risk of going further and selected (also) option c, which stated that the correlation would be higher than 0.5.

With the next research question, *Do students also have a correct intuition?* (1c), we wanted to know whether students, regardless of the presence and strength of their belief in transitivity, had a correct intuition, maybe simultaneously with the misconception, that the correlation between X and Z can be 0 or even lower. With this intention, we classified participants' answers as intuition "present" if they selected any of the correct options in the item (b or d, a for the proportion of variance version) or as intuition "absent" otherwise. We found that 64% of our students (178 out of $n = 279$), including 82% of those who committed the misconception (113 out of $n = 137$), presented no knowledge that the correlation between X and Z can take 0 or negative values because they did not select any of the "present" options.

4.2. RELATION OF STUDENT/TASK-BASED FACTORS AND TRANSITIVITY

In order to address the first research question concerning factors affecting the presence of the misconception, *Are students' characteristics related to the appearance of transitivity?* (2a), we registered the following participants' information along with their responses: age, gender, year at university, number of statistics courses taken at university before the present one, and field of study. Most participants (85%) were between 19 and 23 years old, female (67%), were enrolled in their first (71%) or third (13%) year at university, and had taken one (27%) or no (65%) statistics course before. The distribution among fields of study was as follows: Psychology: 36%, Medicine: 27%, Business: 12%, Statistics: 11%, Mathematics: 11%, Other (or missing data): 3%.

We analyzed the relation of these characteristics with the appearance of the transitivity misconception for both Coding 1 and Coding 2 for the variable transitivity (values "0" and "1" or "0," "1," and "2" depending on the coding) using students' characteristics as independent categorical explanatory variables, but no significant effects were found for any of them.

Concerning the second research question, *Do the type of representation, the size of the correlations, and the contextualization have an effect on the appearance of the transitivity misconception?* (2b), Table 3 shows the frequencies of students showing evidence of the misconception (value "1" for Coding 1 and 2) and the percentages with respect to the total

number of participants answering each type of item. The direction of the effect for the type of representation factor revealed that students who answered the transitivity item in terms of Pearson's correlation coefficient committed the misconception more than students who were confronted with the proportion of variance version of the problem. Respectively, the size of correlations effect showed a higher percentage of students falling into the misconception for the versions of the questionnaire with values 0.5 (0.25 in the proportion of variance version) and 0.7 (0.49 for the proportion of variance version) for the correlation between X and Y , and Y and Z . Finally, the results for the contextualization variable suggested that contextualized problems provoked transitivity less often than theoretically worded items.

Table 3. Presence of transitivity by task factors (coding 1, $n = 271$)

Task factor	Evidence of transitivity misconception present
Type of representation	
Pearson	87 (64.44%, $n = 135$)
Proportion	50 (36.76%, $n = 136$)
Size of correlations	
0.3	36 (39.13%, $n = 92$)
0.5	56 (62.22%, $n = 90$)
0.7	45 (50.56%, $n = 89$)
Contextualization	
Context	63 (45.99%, $n = 137$)
Theoretical	74 (55.22%, $n = 134$)

To test the statistical significance of these effects, a model for the dependent variable transitivity (values "1" or "0" according to Coding 1) was built using type of representation, size of correlations, and contextualization as independent categorical explanatory variables. The results showed statistically significant main effects for the type of representation ($p < 0.0001$) and the size of correlations ($p = 0.0031$), with the contextualization effect on the boundary for statistical significance at the 0.05 level ($p = 0.0525$). None of the interaction effects were statistically significant for any standard level. Based on this finding, we constructed the same model for the variable transitivity according to Coding 1, without the interaction effects, and the results showed again type of representation and size of correlations as the only significant factors (Type: $p < 0.0001$, Size: $p = 0.0049$, Context: $p = 0.1098$). Consequently, we removed the factor contextualization and the resulting model seemed to confirm those results (Type: $p < 0.0001$, Size: $p = 0.0052$) and fit the data well because the Hosmer-Lemeshow goodness-of-fit test was not significant ($df = 4$, Chi-Square = 1.97, $p = 0.7419$).

When similar analyses were performed for Coding 2, results showed that the directions of the effects were as in Table 4.

We see that, in addition to confirming the results for Coding 1, extra information was found. First, in the Pearson condition only 7% of the students provided a completely correct answer as opposed to 21% in the proportion of variance version. Second, when comparing the contextualized and non-contextualized questionnaires, a smaller difference between the two types of item was observed and the percentages answering completely correctly were similar in both versions (15% for the context condition and 13% for the non-contextualized). Finally, for the questionnaires with different sizes of the correlations involved in the transitivity problem it is remarkable that the percentage of students answering completely correct to the item was higher for the condition with large values for the correlations (19%) than for the other two versions (14% for the small correlations

Table 4. Presence of transitivity by task factors (coding 2, $n = 271$)

Task factors	Evidence of transitivity misconception present	Completely correct answer
Type of representation		
Pearson	87 (64.44%, $n = 135$)	9 (6.67%, $n = 135$)
Proportion	50 (36.76%, $n = 136$)	28 (20.59%, $n = 136$)
Size of correlations		
0.3	36 (39.13%, $n = 92$)	13 (14.13%, $n = 92$)
0.5	56 (62.22%, $n = 90$)	7 (7.78%, $n = 90$)
0.7	45 (50.56%, $n = 89$)	17 (19.10%, $n = 89$)
Contextualization		
Context	63 (45.99%, $n = 137$)	20 (14.60%, $n = 137$)
Theoretical	74 (55.22%, $n = 134$)	17 (12.69%, $n = 134$)

condition and 8% for the medium range correlations). As can be seen in Tables 3 and 4, the middle range for the correlations (0.5) seemed to provoke transitivity not only more often than smaller values (0.3), which was consistent with our expectations that the higher the correlation the more prone students would be to commit the misconception, but also more often than higher values (0.7). This last phenomenon was apparently in contradiction with our expectations. However, we performed a post hoc contrast for the appearance of transitivity when using 0.5 versus 0.7 and found that there was no significant difference between these two situations.

When building a categorical model for the three-level response (“0,” “1,” and “2”) variable transitivity defined by Coding 2 using our three task-based factors as independent explanatory variables, we found similar results to those of the model for Coding 1. Again, the factors type of representation and size of correlation were significant for the appearance of the misconception ($p = 0.0013$ and $p = 0.0358$ respectively) and interaction effects were not significant. Hence, we followed the same strategy as before and built first a model considering only the main effects but no interactions. This model confirmed again the significance of type of representation and size of correlation as the only significant factors (Type: $p < 0.0001$, Size: $p = 0.0057$, Context: $p = 0.2751$). Therefore, we next explored a model that considered only the two significant factors and no interaction effects that confirmed these findings (Type: $p < 0.0001$, Size: $p = 0.006$). Moreover, this last model seemed to be appropriate for the data because the Likelihood Ratio goodness-of-fit test was non-significant (LR test: $df = 4$, Chi-Square = 4.74, $p = 0.3151$).

The next research question, *Do those task-based factors have an effect on the strength of the misconception?* (2c), was answered by analyzing the relation of the three selected factors with the strength of the misconception for those students who showed evidence of transitivity ($n = 137$). Table 5 shows that the directions of the effects here revealed that students confronted with the Pearson’s coefficient version of the questionnaires held a soft version of the misconception more often than students confronted with the proportion of variance version. With regard to the size of correlations effect, for those participants completing the 0.7 version (0.49 in the proportion of variance condition), the soft misconception was held less often than for those with the 0.3 (0.09) or 0.5 (0.25) versions. For the contextualization factor there was nearly no effect to be observed.

The first two factors were found again to be statistically significant (Type: $p = 0.0187$, Size: $p = 0.0286$, Context: $p = 0.9910$) when constructing a logistic model for the variable Strength of the Misconception (values “soft” or “strong”). None of the interaction effects were statistically significant for any standard level. Following the strategy above once

Table 5. Strength of transitivity misconception by task factors ($n = 137$)

Task factors	Evidence of soft misconception
Type of representation	
Pearson	61 (70.11%, $n = 87$)
Proportion	25 (50.00%, $n = 50$)
Size of correlations	
0.3	25 (69.44%, $n = 36$)
0.5	40 (71.63%, $n = 56$)
0.7	21 (46.67%, $n = 45$)
Contextualization	
Context	41 (65.08%, $n = 63$)
Theoretical	45 (60.81%, $n = 74$)

more, we constructed first a logistic model without the interaction effects and confirmed the first two factors as the only significant ones (Type: $p = 0.0110$, Size: $p = 0.0129$, Context: $p = 0.4481$). Consequently, we considered next a model for the variable strength using only type of representation and size of correlations as independent categorical explanatory variables. Again, both constructors were significant (Type: $p = 0.0105$ and Size: $p = 0.0153$) and the model seemed to fit the data well because the Hosmer-Lemeshow goodness-of-fit test was non-significant (HL test: $df = 4$, Chi-Square = 0.14, $p = 0.977$).

No significant results were found for the effect of our three task-based factors on the variable intuition (values “present” or “absent”), when addressing the final research question, *Do those factors have an effect on the intuition that the correlation between X and Z can be 0 or smaller?* (2d), by means of logistic modeling.

4.3. CONFIDENCE IN THE ANSWER

In order to deal with the first research question about students’ confidence in transitivity, *What is the average confidence of participants in their answer to the transitivity item as compared to other items?* (3a), we first compared the average confidence selected for this question with the average confidence for the other items in the questionnaire. The average confidence for all students in their answer to the transitivity item was 3.75 on a 0 to 10 scale. When looking at the rest of the questionnaire, this confidence was statistically significantly lower than the average confidence in any other item (see Table 6, the sample sizes correspond to students answering both the specific item and the transitivity item, and the average confidence level for both answers is reported). One exception to this was Item 1.4, for which students showed the lowest confidence (see Appendix for details on this item). Moreover, the average confidence in the answer to the transitivity item was significantly lower than the averaged confidences in the other questions all together ($t = -8.40$, $df = 255$, p -value < 0.0001).

With the intention of addressing the second research question concerning the confidence issue, *Does the degree of confidence in the answer differ significantly between students falling into the misconception and those showing no evidence of it?* (3b), we compared the average confidence in the answer to the transitivity item between students who committed the misconception and students who did not according to Coding 2. This analysis resulted in averages of 4.62 ($n = 37$, $SD = 2.76$) for students providing a completely correct answer, 3.69 ($n = 137$, $SD = 2.44$) for students showing transitivity, and 3.73 ($n = 96$, $SD = 2.93$) for students who committed some other error. Although the differences among these three groups were not significant, the difference

Table 6. Paired *t*-tests for confidence in transitivity item vs. confidence in other items

Item	<i>n</i>	Average confidence in item response	Average confidence in transitivity item response	<i>t</i>	<i>df</i>	<i>P</i>
1.1	125	4.33 (SD = 2.67)	3.62 (SD = 2.39)	-3.46	124	0.0007
1.3	125	5.69 (SD = 2.60)	3.62 (SD = 2.39)	-9.56	124	< 0.0001
1.4	125	2.98 (SD = 2.48)	3.62 (SD = 2.39)	2.86	124	0.0050
1.5/2.2	256	5.07 (SD = 2.58)	3.75 (SD = 2.68)	-7.36	255	< 0.0001
2.3	131	6.15 (SD = 2.60)	3.87 (SD = 2.92)	-8.75	130	< 0.0001
2.4	131	5.58 (SD = 3.02)	3.87 (SD = 2.92)	-5.61	130	< 0.0001
2.5	131	4.74 (SD = 2.81)	3.87 (SD = 2.92)	-3.19	130	0.0018

between those students answering the transitivity item completely correctly and those committing any error (transitivity or other) was on the boundary for statistical significance (t -value = 1.95, df = 268, p = 0.0525).

To answer the third research question related to the confidence of students in their answers, *Are students' characteristics related to the degree of confidence in the answer to the transitivity item?* (3c), we analyzed the relation among participants' confidence in their response to the transitivity item and all registered characteristics (birth date, field of study, year, gender, and statistical background). From all participants' attributes, only gender was statistically significant for the confidence level in the answer to the transitivity item (ANOVA: Gender: df = 1, F -value = 8.05, p = 0.0050). The average confidence of female participants in their answer was 3.38 (SD = 2.48, n = 181), significantly lower than the confidence of male participants (4.80, SD = 2.86, n = 85).

Finally, with regard to the last research question for this study, *Do the three task-based factors have an effect on the degree of confidence in the answer to the transitivity item?* (3d), we found that not one of the three factors was significant when carrying an analysis of variance for the variable confidence. Nevertheless, although not statistically significant, a remarkable phenomenon can be observed when looking at the average confidence of students by factors' levels. Table 7 shows that although participants fell into the misconception more often when confronted with an item stated in terms of Pearson's correlation coefficient, they showed a lower level of confidence for their answers to those problems using the proportion of variance terminology. The same phenomenon could be observed again for the factor size of correlations; participants showed higher confidence when answering problems that deal with the value 0.5 for the correlations between X and Y and Y and Z (0.25 for the proportion of variance version), this being the value that resulted in the transitivity misconception most often.

Table 7. Average confidence by task factors

Task factors	Average confidence
Type of Representation	
Pearson	4.10 (SD = 2.56, n = 135)
Proportion	3.56 (SD = 2.77, n = 135)
Size of Correlations	
0.3	3.74 (SD = 2.84, n = 91)
0.5	4.21 (SD = 2.71, n = 90)
0.7	3.54 (SD = 2.44, n = 89)
Contextualization	
Context	3.83 (SD = 2.50, n = 137)
Theoretical	3.83 (SD = 2.85, n = 133)

5. RESULTS FROM THE INTERVIEW

In order to facilitate the interpretation of the quantitative results obtained from the responses to the questionnaire, we invited all participants to take part in a semi-structured interview with one of the authors. The characteristics of the nine participants in the interviews did not differ much from the complete sample of students participating in the study. This is true for the distribution of gender, age, level of statistical background, and current year at university. The only student-based factor for which participants in the interview differed from the complete sample was their field of study, namely that most interviewees came from a mathematical or statistical environment, as opposed to the high proportion of psychology and medicine students in the complete sample. With regard to the task-based factors, most of our nine participants (randomly) received a contextualized version of the problem (six students) and/or one that presented the correlation in terms of proportion of variance explained by (seven students). The sizes of the correlations in the problems received by interviewees were spread among the three values (four students received the 0.7 version, three student the 0.3, and two students the 0.5).

Most information emerging from the interviews concerned three issues: The motivation of participants about statistics in general, their justification for the selected answers to the items in the questionnaire, and the reasons behind their confidence level in those answers.

First, concerning their general motivation for statistics, most students who volunteered for the interviews confessed to be taking the statistics course only because it was a mandatory requirement for the fulfillment of their program, or because they liked mathematics but found it too difficult so opted for statistics, which they saw as a 'soft and easier' version of mathematics. Just two out of our nine participants in the interview seemed to acknowledge the relevance of statistics for the specific field he or she was involved in.

With regard to the justification of the answers provided by our volunteers to the transitivity item and the reasons for their confidence levels, we repeatedly heard the same arguments during the interviews. More specifically, we could differentiate two types of reasoning behind the answers to the item. First, some students confessed to be just following their intuition based on the context of the question (when present) and their real-life experience. And second, two interviewees explained that they followed "like a formula" or that they tried "all possible calculations" for the two provided correlations to find out about the value of the third correlation, showing a very obvious 'mathematical' reasoning strategy.

When asked why the confidence level in the answer to the transitivity item was lower as compared to most of the other questions (even when those answers were also incorrect), students justified opting for a lower level mainly because of two reasons. They were sure it was not possible to infer the correct answer from the context, or they felt confused about the concepts and words utilized in the question.

Finally, a finding related to the exceptionally low confidence level in question 1.4 was that most of the participants we interviewed were not aware or could not recall learning the relation between the Pearson's correlation coefficient between two variables and the proportion of variance shared by them.

6. DISCUSSION

The present study builds upon the research studies of the last decades that have documented students' misconceptions concerning correlation (see Zieffler, 2006 for a

review). Many of those studies have dealt with younger participants (e.g., Cobb, McClain, & Gravemeijer, 2003; Moritz, 2004). In contrast, our study focuses on university students of introductory statistics courses, as in Batanero and colleagues' line of research (Batanero et al., 1996; Batanero, Estepa, & Godino, 1997; Batanero & Godino, 1998).

The results of the present study confirm the overall presence of the transitivity misconception among our large sample of university students of statistics. Therefore, it justifies the interest in this particular problem for the teaching of introductory statistics and provides an extension for Estepa and Sánchez Cobo's (2001, 2003) classification of students' most common misconceptions about correlation.

Actually, our students not only show evidence of believing that the correlation between X and Z will be positive given positive correlations between X and Y and Y and Z , but also lack a correct intuition that it could be zero (or even negative) in some cases. Nevertheless, more than half of the students in this study committing the misconception merely indicate that the third correlation will be higher than 0, but not higher than 0.5, showing evidence of a so-called soft misconception. This phenomenon of soft misconception accompanying a lack of correct intuition could be due to the fact that this specific property of the correlation was not mentioned during the courses the students had taken, which could explain participants' confusion about this item. When such a specific and exceptional characteristic is not specifically addressed in the statistics classroom, it is very difficult for the students to find out whether their previous knowledge (transitivity of most binary relations) can be extrapolated and in which circumstances it may be misleading.

Concerning our second research interest, it can be concluded from our analyses that expressing the correlation in terms of Pearson's coefficient forces transitivity more than using the proportion of variance representation for our participants. The information extracted from the interviews suggests that students could indeed be more familiar with Pearson's coefficient and hence more prone to deal with it in a routine fashion, extrapolating well-known mathematical properties for binary relations such as transitivity, or trying to apply mathematical formulas.

With regard to the significance of the size of the correlations, we can confirm that, generally speaking, higher values for the two initial correlations provoke more transitivity than smaller values for our sample of students. This phenomenon is in line with the mathematical interval for the correlation between X and Z indicated in the introduction, that shifts to the right the higher the values for the correlations between X and Y , and Y and Z . Furthermore, our data corroborate that the highest level available for the two initial correlations (0.7) is related to the so-called strong version of the misconception more often than the other lower values for our participants.

Finally, regarding the contextualization of the problem, the results are surprising. Our findings provide new and intriguing insight into the role of context in students' problem solving. There is little research available about this topic for the case of statistics, as most of these types of study focus on mathematics problems (e.g., Roth, 1997). In our study, the context for the item was written with the intention of creating a prone-to-transitivity context. It seemed reasonable to expect that students believed in transitivity for the topic of school subject grades, as it is common for them to find that "good" students tend to have higher scores in general as opposed to "worse" students who tend to have lower scores in all subjects, therefore confirming correlations between subject scores. However, the context seems to actually help participants avoid the misconception and theoretically written items motivate the misconception more often. Despite these results and the evidence found during the interviews confirming the relevance of context for the reasoning behind participants' answers, the effect of this task-based factor is not

statistically significant, and hence more research is needed to explore the effect of (eventually different) context(s) on the appearance of the misconception.

Concerning the third and last research aim for our study, the investigation of the confidence of participants in their answers to the items, our findings contribute to the literature that examines confidence as a central factor affecting academic proficiency (Stajkovic & Luthans, 1998), specifically for statistics courses (Finney & Schraw, 2003). We found sparse previous research in the statistics education literature about the relation between confidence in a specific concept and the correctness of such a concept; and those studies that looked into it have reported mixed results. For example, Allen, Reed Rhoads, and Terry (2006) found a positive trend between correct responses and confidence across items in their study about introductory statistics concepts. On the other hand, Chance, delMas, and Garfield (2004) did not find any relevant relation between students' confidence in their responses to items about sampling distributions and the correctness of those answers. Our results extend this evidence but cannot confirm a strong connection either. In fact, the difference in confidence between students providing a completely correct answer to the transitivity item and students providing an incorrect (misconception or other) answer is in the boundary of being statistically significant for our study. Our findings reveal as well that students taking part in this study answer the question about the transitivity property of the correlation with lower confidence than almost all other items in the questionnaire. One exception is the item about the relation between the correlation coefficient and the proportion of variance, to which they respond with even less confidence. These results show participants are confused about the link between the correlation coefficient and the proportion of variance explained, and are uncertain when extracting conclusions from two correlations among three variables.

Finally, interesting results are found when exploring the effects of task-based factors on the confidence of participants. Again, the first two factors (type of representation and size of correlations) have a significant effect on the confidence of our students and all three (including contextualization) reveal an unexpected direction in their effects as compared to their effects on the correctness of the answer. In other words, those levels of the factors that provoke more transitivity (Pearson's representation, correlations of 0.5, theoretical wording) are related to higher levels of confidence. One possible conclusion seems to be in line with the reactions observed during the interviews. Namely, that participants are more confident in situations where they can follow a previously learned rule such as transitivity, therefore committing the misconception, than when they are forced to use independent reasoning, which provokes the misconception less often but makes them feel more insecure about their answers.

7. CONCLUSION

Research about misconceptions is of paramount importance for helping students learn statistics more effectively (Rossman & Chance, 2004). Specifically, new insights into correlational reasoning are relevant because it is one of the most important cognitive activities humans perform (McKenzie & Mikkelsen, 2007) and central within the introductory statistics course (Garfield, 2003). With the present study, we contribute to such research by expanding the available knowledge base of misconceptions about correlation. We provide evidence of a new misconception, the transitivity misconception of the correlation, and investigate student- and task-based factors related to it. In addition, we report on the relation between the presence of the misconception and individual task-specific confidence.

The main conclusion that can be extracted from this study is that the transitivity misconception of the correlation seems to be widespread among our students of introductory courses of statistics and that it could be added to the list of common students' misconceptions about correlation suggested by Estepa and Sánchez Cobo (2001, 2003). If students are expected to be able to understand and carry out correlational studies after a university statistics course, instructors may be in position to help remedy students' misconceptions in this area. One possibility would be presenting students with examples of real-life situations where the non-transitivity becomes evident. This is in line with the recommendations from several authors to confront students with their misconceptions as an effective measure for enhancing learning (Garfield, 1995; Lipson, 2002; Pfannkuch & Brouwn, 1996; Pressley & McCormick, 1995; Resnick, 1987; Smith III, diSessa, & Roschelle, 1993). Emphasis on the importance of being able to reason about correlations in an independent manner regardless of task-based factors such as contextualization and type of representation as a first step before interpretation could be a focus for teachers of introductory statistics.

Our results also clearly point to the need of additional research about the transitivity misconception. First, the validity of the effects of the type of representation and size of correlations factors could be confirmed by means of replications and extensions of the item used here. Second, the contextualization factor should be further explored in order to better understand how students deal with statistical contexts. In fact, there is a general call for further research about the role of context in statistics education, being the topic for the next *International Conference on Teaching Statistics* (ICOTS-8, 2010) "Data and context in statistics education: Towards an evidence-based society." To build upon our study, different types of context (more and less prone to the misconception) could be used and more interviews could be carried out in order to clarify the real influence of this factor or the illusion of students about the factors they actually take into consideration when reasoning about transitivity and correlation. Furthermore, open-answer items would help gain new insights into the presence of and the factors affecting this misconception, as multiple-choice items generally have not been able to provide enough information about the reasoning process of the students (Gal & Garfield, 1997).

In addition to these studies focusing on students of introductory statistics courses, research that explores both teachers' and researchers' understanding of the non-transitivity of the correlation could provide much extra information for the understanding of the situation among students and for finding possible advice to be followed by the teachers to address the misconception in their classrooms.

Finally, once enough information has been gathered, intervention studies could be carried out to analyze the effectiveness of the recommendations mentioned above and others derived from the results from the additional studies.

Overcoming this (and other) misconceptions of correlation would have a great impact on students' proper development and power of interpretation of correlation analyses within their courses and in their prospective professional practices.

ACKNOWLEDGEMENTS

A small part of the results described in this article have been presented at the 56th Session of the International Statistical Institute. This study is part of, and was funded by, research GOA 3H040514 from the Katholieke Universiteit Leuven. The authors would also like to thank the student Virginie März for her help with the analyses of the empirical data, and all teachers and students of the Complutense University (Madrid, Spain) who voluntarily participated in the study.

REFERENCES

- Allen, K., Reed Rhoads, T., & Terry, R. (2006, October). *Work in progress: Assessing student confidence of introductory statistics concepts*. Paper presented at the Frontiers in Education Conference, 36th Annual, San Diego, CA.
- Batanero, C., Estepa, A., & Godino, J. D. (1997). Evolution of students' understanding of statistical association in a computer based teaching environment. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 191-205). Voorburg, The Netherlands: International Association for Statistical Education and International Statistical Institute.
[Online: <http://www.stat.auckland.ac.nz/~iase/publications/8/15.Batanero.pdf>]
- Batanero, C., Estepa, A., Godino, J. D., & Green, D. R. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education*, 27, 151-169.
- Batanero, C., & Godino, J. D. (1998). Understanding graphical and numerical representations of statistical association in a computer environment. In L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee, & W-K. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching Statistics* (pp. 1017-1024). Singapore. Voorburg, The Netherlands: International Statistical Institute.
[Online: <http://www.stat.auckland.ac.nz/~iase/publications/2/Topic7zf.pdf>]
- Beyth-Marom, R. (1982). Perception of correlation reexamined. *Memory and Cognition*, 10, 511-519.
- Chance, B., delMas, R., & Garfield, J. (2004). Reasoning about sampling distributions. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 295-323). The Netherlands: Kluwer Academic Publishers.
- Cobb, P., McClain, K., & Gravemeijer, K. P. E. (2003). Learning about statistical covariation. *Cognition and Instruction*, 21(1), 1-78.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Estepa, A., & Sánchez Cobo, F. T. (2001). Empirical research on the understanding of association and implications for the training of researchers. In C. Batanero (Ed.), *Training researchers in the use of statistics* (pp. 37-51). Granada, Spain: International Association for Statistical Education and International Statistical Institute.
- Estepa, A., & Sánchez Cobo, F. T. (2003). Evaluación de la comprensión de la correlación y regresión a partir de la resolución de problemas [Evaluation of the understanding of correlation and regression through problem solving]. *Statistics Education Research Journal*, 2(1), 54-68.
[Online: [http://www.stat.auckland.ac.nz/~iase/serj/SERJ2\(1\).pdf](http://www.stat.auckland.ac.nz/~iase/serj/SERJ2(1).pdf)]
- Finney, S. J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology*, 28, 161-186.
- Gal, I. & Garfield, J. B. (1997). *The assessment challenge in statistics education*. Amsterdam, The Netherlands: IOS Press and the International Statistical Institute.
- Garfield, J. (1995). How students learn statistics. *International Statistical Review*, 63, 25-34.
- Garfield, J. B. (2003). Assessing statistical reasoning. *Statistics Education Research Journal*, 2(1), 22-38.
[Online: [http://www.stat.auckland.ac.nz/~iase/serj/SERJ2\(1\).pdf](http://www.stat.auckland.ac.nz/~iase/serj/SERJ2(1).pdf)]

- Inhelder, B. & Piaget, J. (1955). *De la logique de l'enfant à la logique de l'adolescent* [From the logic of the child to the logic of the adolescent]. Paris: Presses Universitaires de France.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, MA: Cambridge University Press.
- Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction*, 6, 59-98.
- Konold, C. (1991). Understanding students' beliefs about probability. In E. Von Glaserfeld (Ed.), *Radical constructivism in mathematics education* (pp. 139-156). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Langford, E., Schwertman, N., & Owens, M. (2001). Is the property of being positively correlated transitive? *The American Statistician*, 55(4), 322-325.
- Lipson, K. (2002). The role of computer based technology in developing understanding of the concept of sampling distribution. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics: Developing a statistically literate society*, Cape Town, South Africa. [CD-ROM]. Voorburg, The Netherlands: International Statistical Institute.
[Online: http://www.stat.auckland.ac.nz/~iase/publications/1/6c1_lips.pdf]
- McKenzie, C. R. M., & Mikkelsen, L. A. (2007). A Bayesian view of covariation assessment. *Cognitive Psychology*, 54, 33-61.
- McNemar, Q. (1969). *Psychological statistics* (4th ed.). New York: John Wiley and Sons, Inc.
- Moore, T. L. (2006). Paradoxes in film ratings. *Journal of Statistics Education*, 14(1).
[Online: <http://www.amstat.org/publications/jse/v14n1/datasets.moore.html>]
- Moritz, J. B. (2004). Reasoning about covariation. In D. Ben-Zvi & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 227-256). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice Hall.
- Pfannkuch, M., & Brouwn, C. M. (1996). Building on and challenging students' intuitions about probability: Can we improve undergraduate learning? *Journal of Statistics Education*, 4(1).
[Online: <http://www.amstat.org/publications/jse/v4n1/pfannkuch.html>]
- Pressley, M., & McCormick, C. (1995). *Advanced educational psychology for educators, researchers, and policymakers*. New York: Harper-Collins.
- Resnick, L. B. (1987). Learning in school and out. *Educational Researcher*, 16(9), 13-20.
- Rossmann, A. J., & Chance, B. (2004, August). *Anticipating and addressing student misconceptions*. Paper presented at ARTIST Roundtable Conference, Appleton, WI
[Online: <http://www.rossmanchance.com/artist/proceedings/rossman.pdf>]
- Roth, W.-M. (1997). Where is the context in contextual word problems? Mathematical practices and products in grade 8 students' answers to story problems. *Cognition and Instruction*, 14, 487-527.
- Shaughnessy, J. M. (2003). Research on students' understanding of probability. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 216-226). Reston, VA: National Council of Teachers of Mathematics.
- Smith III, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences*, 3(2), 115-163.

- Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin*, *124*(2), 240-261.
- Zieffler, A. (2006). *A longitudinal investigation of the development of college students' reasoning about bivariate data during an introductory statistics course*. Unpublished doctoral dissertation, University of Minnesota, USA.

ANA ELISA CASTRO SOTOS
Centre for Methodology of Educational Research
Katholieke Universiteit Leuven
Andreas Vesaliusstraat 2
3000 Leuven
Belgium

APPENDIX

Item 1.1 about the commutative property of the correlation:

Suppose that the correlation between two variables, grade in mathematics and attitudes toward mathematics, is 0.78. What does that mean about the relation between the attitudes toward mathematics and the grade in mathematics?

- a. *Nothing, there is not enough information*
- b. *The correlation is -0.78*
- c. *The correlation is $1-0.78=0.22$*
- d. *The correlation is still 0.78*

How confident are you that your answer will be correct?

0 1 2 3 4 5 6 7 8 9 10

No confidence 100% confidence

Item 1.2 and item 2.1 were the transitivity item.

Item 1.3 about the causality misconception:

We find that the correlation between Educational level and income is 0.72. What does this mean?

- a. *Higher income results from a higher Educational level*
- b. *A low income is related to a higher Educational level*
- c. *People with a lower Educational level tend to have lower income*
- d. *People with a higher Educational level tend to have lower income*

How confident are you that your answer will be correct?

0 1 2 3 4 5 6 7 8 9 10

No confidence 100% confidence

Item 1.4 about the relation between the correlation coefficient and the proportion of variance:

Which of the following numbers will be higher, the correlation between X and Y, or the proportion of variance of Y explained by X? (Assume a positive correlation between X and Y.)

- a. *They will be the same*
- b. *The proportion of variance of Y explained by X will be higher because it is the square of the correlation*
- c. *The proportion of the variance of Y explained by X will be lower because it is the square of the correlation*
- d. *It is not possible to know, these two indexes are not related*

How confident are you that your answer will be correct?

0 1 2 3 4 5 6 7 8 9 10
 No confidence 100% confidence

Item 1.5 and item 2.2 about the causality misconception:

During a month, 500 students from a certain school registered the daily hours they spent watching television. The weekly average amount of hours dedicated to watch television was 28. The researchers who were responsible for the study also looked at the school reports for each student. They discovered that those students who obtained good school results dedicated less time to watch television than those students who obtained mediocre results. Below we list several possible conclusions about the results of this investigation. Select all conclusions that you agree with.

- a. The sample of 500 is too small to allow obtaining conclusions
- b. If a student diminished the time dedicated to watch television, his/her performance would improve
- c. Even though the best students watched less television, this does not necessarily imply that watching television harms school performance
- d. A month is not long enough to estimate how many hours students actually dedicate to watch television
- e. The investigation demonstrated that watching television weakens school performance
- f. I do not agree with any of these conclusions

How confident are you that your answer will be correct?

0 1 2 3 4 5 6 7 8 9 10
 No confidence 100% confidence

Item 2.3 about the definition of hypotheses tests:

Select the definition of "hypothesis test/contrast" that you consider most correct:

- a. Proof of the truth or falseness of the null hypothesis
- b. Proof of the truth or falseness of the alternative hypothesis
- c. Proof of the probability or improbability of the null hypothesis
- d. Proof of the probability or improbability of the alternative hypothesis
- e. Assessment of the evidence in the data in favor of or against the null hypothesis
- f. Assessment of the evidence in the data in favor of or against the alternative hypothesis

How confident are you that your answer will be correct?

0 1 2 3 4 5 6 7 8 9 10
 No confidence 100% confidence

Item 2.4 about the interpretation of a p -value:

