LEARNING PROBABILITY AND STATISTICS: COGNITIVE AND NON-COGNITIVE FACTORS RELATED TO PSYCHOLOGY STUDENTS'ACHIEVEMENT

Francesca Chiesi and Caterina Primi
Department of Psychology, University of Florence, Italy
francesca.chiesi@unifi.it

The aim of the present is to ascertain the impact of both cognitive and non-cognitive factors on probabilistic and statistics reasoning in psychology students enrolled in introductory statistics courses. It was hypothesised that performance was related to the student's general and mathematical background (cognitive factors), math self-efficacy and attitudes toward statistics (non-cognitive factors). A structural equation model was specified in which cognitive and non-cognitive factors were considered as the exogenous latent variables having an impact on both probabilistic and statistics reasoning. Results stressed the role of both cognitive and non cognitive factors suggesting that competence as well as attitudes and self-efficacy should be the focus in planning interventions to help students in increasing performance.

INTRODUCTION

Many students find it difficult to grasp probability and statistical concepts, as documented in different educational contexts, and it seems especially true for students attending graduate programmes that are traditionally qualitative, as degrees in Psychology, Education, Health Sciences. For this reason, more attention has been paid to psychology student's beliefs and feelings about statistics, and several researches have focused on the identification of models with non-cognitive factors such as beliefs and feeling about statistics in order to better understand the underlying mechanism of statistics achievement (Zieffler et al., 2008).

Tremblay and colleagues (2000) started from Lalonde and Gardner's (1993) study in which statistics learning was conceptualized as similar to second language learning. Lalonde and Gardner tested a model with psychology students in which achievement was related to individuals' mathematical aptitude, statistics anxiety and attitudes as well as motivation to learn statistics and effort. Analyses revealed that all the hypothesized paths were significant except the direct path between achievement and anxiety: performance in statistics was directly predicted by math aptitude and motivation, whereas attitudes and anxiety influenced indirectly achievement through motivation. Introducing some indicators for academic achievement, Tremblay et al. (2000) replicated Lalonde and Gardner's results and, in contrast to the previous study, they found that anxiety also affected achievement directly. Furthermore, attitudes had an effect on anxiety: unfavourable attitudes toward the course resulted in high levels of anxiety which, in turn, reduced performance.

Onwuegbuzie (2003) proposed a model of statistics achievement amongst graduate Psychology and Educational disciplines, called the *Anxiety-Expectation Mediation* (AEM) model. As it could be argued that learning statistics is akin to learning a foreign language (Lalonde & Gardner, 1993), the same model was put forward for statistics achievement. Results confirmed the link between anxiety and performance (Tremblay et al., 2000; Onwuegbuzie & Seaman, 1995; Zeidner, 1991). The relationship between performance and expectation on own performance is considered as an important manifestation of self-efficacy (Bandalos et al., 1999), so anxiety and expectation can be considered the best predictors of achievement in statistics courses.

Nasser (2004) examined the extent to which anxiety and attitudes toward mathematics and statistics, motivation and mathematical aptitude explained the achievement of pre-service teachers enrolled in an introductory statistics course. They measured all factors considered to be related to achievement during the course before the midterm examination. They reported a high positive effect of mathematical aptitude and a lower, but significant, positive effect of attitudes on performance. Mathematics anxiety was found to be directly linked to attitudes: a strong negative effect was reported indicating high level of mathematics anxiety related to low level of positive attitudes toward statistics. More recently, Budé et al. (2007) focused on motivational constructs (partially overlapping with some dimensions of attitude towards statistics) finding that affect towards the discipline influenced performance strongly in Health Sciences students. Finally,

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

Dempster and McCorry (2009) examine the nature of the relationships between undergraduate psychology students' previous experiences of maths, statistics and computing, their attitudes toward statistics, and assessment on a statistics course. Of the variables examined, the strongest predictor of assessment outcome was students' attitude about their intellectual knowledge and skills in relation to statistics.

The literature review suggests that the relationships between cognitive and non-cognitive factors related to statistics achievement are quite complex. Starting from these premises, the present study is aimed at better ascertaining the impact of both cognitive and non-cognitive factors on probabilistic and statistics reasoning in psychology students enrolled in introductory statistics courses. Teaching statistics with psychology students, i.e. with students who are not primarily interested in statistics, produces in difficulties (Wiberg, 2009), and it turns out to be relevant examining the role of students' knowledge, beliefs about their own ability in dealing with statistics, and their feeling about the topic. Indeed, students entering these disciplines sometimes do not have a strong mathematics background, often are not confident about their capacities, and dislike anything "mathematical" (Dempster & Mc Corry, 2009).

More specifically, it was hypothesised that achievement was related directly to cognitive factors measured by the student's general background and mathematical competence, and non-cognitive factors identified by the mathematics self-efficacy and the attitudes toward statistics concerning students' beliefs about their own ability and feeling toward the discipline. Recently, Konold and Kazak (2008) suggested that some of the difficulties students have in learning basic skills in data analysis stem from a lack of rudimentary idea in probability. Starting from this assumption, achievement was operationalized including both probabilistic and statistics reasoning. The causal paths among variables related to achievement are explored using structural equation modelling techniques in which cognitive and non-cognitive factors were considered as the exogenous latent variables having an impact on achievement, the endogenous latent variable in the model.

METHODS

Participants

Data were collected from 113 psychology students attending the University of Florence in Italy enrolled in undergraduate introductory statistics course. The course was scheduled to take place over 10 weeks, and takes 6 hours per week (for a total amount of 60 hours). It covered the usual introductory topics of descriptive and inferential statistics (including basic concept of probability theory and data analysis), and their application in psychological research. Participants mean age was of 21.6 years (SD = 5.13), and most of the participants were women (87%). This proportion reflects the gender distribution of the population of psychology students in Italy. All students participated on a voluntary basis after they were given information about the general aim of the investigation (i.e., collecting information in order to improve students' statistics achievement).

Measures

General Background Test (GBT). This is a scholastic assessment test consisting in 100 multiple choice questions (one correct out of five alternatives) divided in five sections: mathematics, biology, English comprehension, critical reading and reasoning. The time for the test was 85 minutes. A single composite score, based on the sum of correct answers less the wrong answers (the score for a wrong answer was -.25) was calculated.

Prerequisiti di Matematica per la Psicometria (PMP, Galli, Chiesi & Primi, 2008). In order to develop a scale to measure accurately the mathematics ability needed by psychology students enrolling in introductory statistics courses, the PMP scale was constructed applying the Rasch model, and its reliability and validity was tested (Galli et al., 2008; Ciancaleoni, et al., 2008). The contents were defined on the basis of the basic mathemathics abilities required to solve descriptive and inferential statistics problems. The PMP scale is a 30-problem test. Each problem presents a multiple choice question (one correct out of four alternatives). A single composite, based on the sum of correct answers, was calculated.

Survey of Attitudes Toward Statistics (SATS) (Schau et al., 1995). The SATS provides a multidimensional measure of attitude that includes the perception of statistics in itself and as part of the degree programme, as well as affective and cognitive components. In this work we administered the Italian version of SATS previously validated (Chiesi & Primi, 2009).The SATS contains 28 Likert-type items using a 7-point scale ranging from *strongly disagree* to *strongly agree*. The SATS assesses four Attitudes components: *Affect* measures positive and negative feelings concerning statistics; *Cognitive Competence* measures students' Attitudes about their intellectual knowledge and skills when applied to statistics; *Value* measures Attitudes about the usefulness, relevance, and worth of statistics in personal and professional life; *Difficulty* measures students' Attitudes about the difficulty of statistics as a subject.

Mathematics Self-Efficacy Scale revised (MSES-R) (Kranzler & Pajares, 1997; Pajares & Miller, 1995). The MSES-R asks students to express their level of confidence in successfully solving mathematics problems. The English version of this scale is widely employed with high school students and freshmen university students. This scale has been developed to assess the students' confidence to solve math problems (*Solution of Mathematics Problems*), everyday math tasks (*Completion of Mathematics Tasks*), and to attain a satisfactory performance in college courses (*Performance in College Mathematics Courses*). In this work we administered the Italian version of the subscale *Solution of Math Problems*, whose psychometric properties had been previously assessed (Chiesi, Primi & Galli, 2009). The subscale is composed by 18 math problems with different difficult levels equally allocated to three areas (arithmetic, algebra and geometry).

Probabilistic Reasoning Questionnaire (PRQ) (Chiesi, Primi, & Morsanyi, 2009). The questionnaire contained 10 multiple-choice probabilistic reasoning tasks (one correct out of three alternatives). Each task was scored either 1 (correct) or 0 (incorrect) based on whether the participant gave the normatively appropriate response for the task. The scores on the probabilistic reasoning tasks were summed to form a composite score.

Introductory Statistics Inventory (ISI) (Chiorri, Piattino, Primi, Chiesi, & Galli, 2009). The ISI is a test consisting of 30 multiple-choice items (one correct out of four alternatives) to evaluate achievement after an introductory Statistics course. Half problems are descriptive and other half inferential. Each task was scored either 1 (correct) or 0 (incorrect) and a composite score was obtained.

Procedure

The GBT was administered before the beginning of the course. The SATS was presented at the beginning of the course during the first day. The PMP was completed during the second day of class, and the PRQ and the ISI at the end of the class.

RESULTS

Significant correlations were found between the measured variables (Table 1) that support the hypothesised relationships.

Table 1. Means, standard deviations (in brackets), and correlations among the measured variables

Variables	*M*	*s*	1	2	3	4	5	6
1. General Background	*44.43*	*10.32*						
2. Maths Basics	*23.10*	*4.20*	.29*					
3. Maths Self-efficacy	*83.53*	*12.52*	.32**	.22**				
4. Cognitive Competence	*26.73*	*6.19*	.18	.15	.41**			
5. Affect	*22.79*	*6.87*	.11	.11	.29*	.76**		
6. Probabilistic Reasoning	*6.29*	*1.50*	.40**	.22**	. 34*	.25*	.23*	
7. Statistic Reasoning	*20.01*	*4.72*	.39**	.34**	.26*	.23*	.19*	.25*

* $p < .05$
** $p < .01$

The model included 3 latent variables and 9 manifest variables (Figure 1): Cognitive Factors and Non Cognitive Factors were the exogenous latent variable that influenced directly the Statistic and Probabilistic Achievement. Cognitive Factors were measured through the the GBT (General Background) and the PMP (Maths Basics). The two scores of the SATS subscales (Cognitive Competence and Affect) and the score of the MSES-R (Maths Self-Efficacy) were used as indicators of Non Cognitive Factors. A covariance path was traced between errors of subscales measuring the attitude dimensions. The endogenous latent variable (Probability and Statistics Achievement) was measured through two scores of PRQ subscales - obtained by dividing the test randomly in two parts with 5 items for each (Probability 1 and Probability 2) - and two scores of PRQ subscales - obtained considering separately the descriptive and the inferential items (Descriptive and Inferential). Covariance paths were traced between the two indicators derived from the PRQ and the two indicators derived from the ISI.

SEM analyses were conducted with AMOS 5.0 (Arbuckle, 2003) using maximum likelihood estimation on the variance-covariance matrix. Univariate distributions of all variables included in the model and their multivariate distribution were examined for assessment of normality. Skewness and curtosis indices (ranging respectively from -.44 to .19, and -.87 to .77) attested that the departures cannot be expected to lead to appreciable distortions (Marcoulides & Hershberger, 1997; Muthén & Kaplan, 1985) with the exception of the PMP score distribution that was negatively skewed (*Skewness* = -1.21). A log-transformation ($\text{Log}_{10}\,(K - X)$, where $K = X_{\max} + 1$] was conducted in order to reduce the skewness. The index of Multivariate Kurtosis (Mardia, 1970) (β=2.57, *c.r.*=0.97, *p*=.17) indicated that there was not a significant departure from multivariate normality. That is, data met the assumption of multivariate normal distribution required by SEM.

The model showed a good fit to data ($\chi^2(23)=32.93$, $p=.08$; χ^2/d.f.= 1.43; RMSEA= .062; CFI= .96; NNFI= .94), and all the estimated structural coefficients were statistically significant (Figure 1). As expected Cognitive and Non Cognitive factors had a significant direct effect on Probability and Statistics Achievement. However, the relation with the Cognitive factors was more strong (.85 *vs* .48).

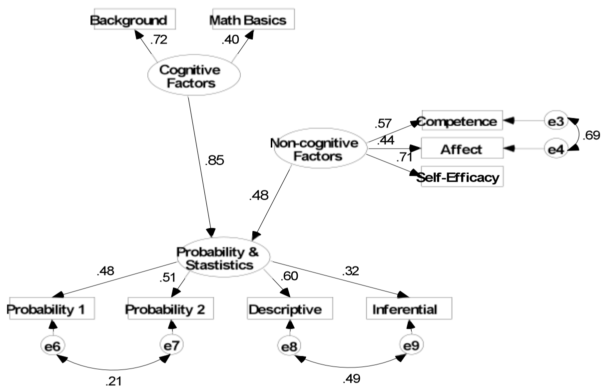

Figure 1. Model of Statistics Achievement with standardized parameters
(paths are all significant at the .05 level or lower)

CONCLUSION

The purpose of the present research was to examine psychology students' achievement in introductory statistics course in order to better ascertain the impact of both cognitive and non-cognitive factors on course performance. As expected, and in line with previous research (Lalonde & Gardner, 1993; Tremblay et al., 2000), mathematical knowledge, acquired during high school, had a direct and strong effect on achievement. Additionally, mathematics self-efficacy and attitudes toward the discipline affected achievement (Dempster & McCorry, 2009), that is, perceived competences and affect concurred in determining performance in statistics.

Given that cognitive and non-cognitive factors concur to determine achievement, there is potential for developing interventions which will modify both competences and perceived competences. Such interventions should focus on helping students to believe that they have the intellectual capacity to cope with the demands of the introductive statistics course, that is not a math course even if it requires some math basics to employ in data analysis. So, it might be important to provide students with some remedial teaching of basic mathematics in order to improve also students' confidence in approaching the subject of statistics.

More in detail, it might be useful to arrange a short series of lessons aimed at mastering the basic mathematical skills (rounding, factorials, order of operations in computations, basic algebra, and so on) required during the course. Indeed, they constitute a necessary tool to keep in touch with statistics whereas the assessment of statistics achievement does not depend solely on these basic mathematical techniques. Moreover, to help students enhance their confidence in learning statistics it might be useful to give exercises that allow students to experience that they can master the topics, and to provide feedback about their results in order to allow them to monitor their progress.

Future research should aim to test the effectiveness of these pedagogical techniques: the impact of interventions directed in enhancing basic mathematical skills should be specifically assessed as well as the related effects on attitudes and performance. Collecting repeated measures of both cognitive and non-cognitive factors during the course might help in monitoring these changes and enhancements.

REFERENCES

Arbuckle, J. L. (2003). *Amos 5.0* [computer software]. Chicago: Smallwaters.

Bandalos, D. L., Finney, S. J., & Geske, J. A. (2003). A model of statistics performance based on achievement goal theory. *Journal of Educational Psychology, 95*(3), 604-616.

Budè, L., Van De Wiel, M. W. J., Imbos, T., Candel, M. J. J. M., Broers, N. J., & Berger, M. P. F. (2007). Students' achievements in a statistics course in relation to motivational aspects and study behaviour. *Statistics Education Research Journal, 6*(1), 5-21. Online: www.stat.auckland.ac.nz/~iase/serj/SERJ6(1).

Chiesi, F., & Primi, C. (2009). Assessing Statistics Attitudes Among College Students: Psychometric Properties Of The Italian Version Of The *Survey Of Attitudes Toward Statistics* (*SATS*): *Learning and Individual Differences, 19*, 309-313.

Chiesi, F., Primi, C., & Galli, S. (2009). *Competenze o autoefficacia in matematica? Uno studio sul rendimento a Psicometria.* Atti del III° Convegno "Didattica e Integrazione del sapere psicologico", 31-43 [documento WWW]. Online: http:\\convdidattica.psy.unipd.it\

Chiesi, F., Primi, C., & Morsanyi, K. (2009). The Effects of Education, Instructions and Cognitive Abilities on Probabilistic Reasoning: A Test of a Theory. *Society for Research in Child Development, SRCD, Biennial Meeting*, Denver, 2-4 Aprile.

Chiorri, C., Piattino, S., Primi, C., Chiesi, F., & Galli, S. (2009). *Why Psychology Students don't Like Statistics?* MathPsych, Amsterdam, 1-4 Agosto.

Ciancaleoni, M., Galli, S., Chiesi, F., & Primi, C. (2008). *Assessing the predictive validity of a mathematical ability scale constructed applying the Rasch model*. III European Congress of Methodology, Oviedo, 8-12 Luglio.

Dempster, M., & McCorry, N.K. (2009). The Role of Previous Experience and Attitudes Toward Statistics in Statistics Assessment Outcomes among Undergraduate Psychology Students. *Journal of Statistics Education* Volume, *17*(2). Online: www.amstat.org/publications/jse/v17n2/dempster.html.

Galli, S., Chiesi, F., & Primi, C. (2008). The construction of a scale to measure mathematical ability in psychology students: an application of the Rasch Model. *TPM - Testing Psychometrics, and Methodology in Applied psychology, 15,* 1-16.

Kranzler, J., & Pajares, F. (1997). An exploratory factor analysis of the Mathematics Self-Efficacy Scale-Revised, MSES-R. *Measurement and Evaluation in Counseling and Development, 29,* 215-228.

Konold, C., & Kazak, S. (2008). Reconnecting Data and Chance. *Technology Innovations in Statistics Education, 2*(1). Online: http://repositories.cdlib.org/uclastat/cts/tise/vol2/iss1/art1/.

Lalonde, R. N., & Gardner, R. C. (1993). Statistics as a second language? A model for predicting performance in psychology students. *Canadian Journal of Behavioural Science, 251*, 108-125.
Mardia, K.V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*, 519–530.
Marcoulides, G. A., & Hershberger, S. L. (1997). *Multivariate statistical methods. A first course*. Mahawa, NJ: Lawrence Erlbaum Associates.
Muthén, B., & Kaplan, D. (1985). A comparison of Some Methodologies for the Factor Analysis of Non-normal Likert Variables. *British Journal of Mathematical and Statistical Psychology, 38*, 171-189.
Nasser, F. (2004). Structural Model of the Effects of Cognitive and Affective Factors on the Achievement of Arabic-Speaking Pre-service Teachers in Introductory Statistics. *Journal of Statistics Education, 12*. Online: www.amstat.org/publications/ jse/v12n1/nasser.html.
Onwuegbuzie, J. A. (2003). Modeling statistics achievement among graduate students. *Educational and Psychological Measurement*, *63*, 1020-1038.
Onwuegbuzie, A. J., & Seaman, M. (1995). The effect of time constraints and statistics test anxiety on test performance in a statistics course. *Journal of Experimental Education, 63*, 115-124.
Pajares, F., & Miller, M. D. (1995). Mathematics self-efficacy and mathematics outcomes: The need for specificity of assessment. *Journal of Counseling Psychology, 42*, 190-198.
Schau, C., Stevens, J. J., Dauphine, T. L., & Del Vecchio, A. (1995). The development and validation of the survey of attitudes toward statistics. *Educational and Psychological Measurement, 55,* 868-875.
Tremblay, P. F., Gardner, R. C., & Heipel, G. (2000). A model of the relationships among measures of affect, aptitude, and performance in introductory statistics. *Canadian Journal of Behavioral Science, 32*, 40-48.
Wiberg, M. (2009). Teaching Statistics in Integration with Psychology. *Journal of Statistics Education*, *17*(1). www.amstat.org/publications/jse/v17n1/wiberg.html
Zeidner, M. (1991). Statistics and mathematics anxiety in social science students: some interesting parallels. *British Journal of Educational Psychology, 61*, 319-328.
Zieffler, A., Garfield, J., Alt, S., Dupuis, D., Holleque, K., & Chang, B. (2008). What Does Research Suggest About the Teaching and Learning of Introductory Statistics at the College Level? A Review of the Literature. *Journal of Statistics Education, 16*(2). Online: www.amstat.org/publications/jse/v16n2/zieffler.html.