Psicothema

Psicothema 2013, Vol. 25, No. 4, 494-499
doi: 10.7334/psicothema2013.15

ISSN 0214 - 9915 CODEN PSOTEG

www.psicothema.com

# Distortions and gender-related differences in the perception of mechanical engineering in high school students

Pere Joan Ferrando Piera, Mar Gutiérrez-Colón Plana, Paloma Paleo Cageao, Silvia de la Flor López and Francesc Ferrando Piera
Universidad Rovira i Virgili

## Abstract

**Background:** The reason for this study was the low interest that high school students, particularly females, show for the subject of mechanical engineering (ME). We assumed that this problem was partly due to: (a) lack of understanding of the tasks involved in ME, and (b) a distorted and negative perception of the professional environment and working conditions. **Method:** To assess these two assumptions, two measurement instruments (tasks and perceptions) were developed and administered in a sample of 496 high school students. A multiple-group design was used and data was analyzed by using an extended item response theory model. **Results:** In general terms, the results agreed with our expectations. However, no significant gender differences were found. **Discussion:** The implications of the results for future improvements are discussed.

***Keywords:*** Mechanical engineering and gender differences, attitude measurement, multi-group factor analysis, item response theory.

## Resumen

***Distorsiones y diferencias de género en la percepción de la ingeniería mecánica en alumnos de Bachillerato.*** **Antecedentes:** la motivación del presente estudio es la baja demanda de la carrera de ingeniería mecánica por parte de los alumnos, y muy especialmente las alumnas, que ingresan en la Universidad. Se planteó que, en parte, esta situación se debía a: (a) un conocimiento erróneo de las tareas propias de la profesión, y (b) una percepción distorsionada y negativa de su entorno y condiciones de trabajo. **Método:** para evaluar estos dos puntos se desarrollaron dos instrumentos de medida (tareas y percepciones) que se administraron en una muestra de 496 alumnos de Bachillerato. Se utilizó un diseño multi-grupo y los datos se analizaron de acuerdo a un modelo extendido de la teoría de respuesta al ítem. **Resultados:** en términos generales los resultados apoyan los supuestos de partida. Sin embargo, no se encontraron diferencias significativas entre hombres y mujeres. **Discusión:** finalmente, se discuten las implicaciones de los resultados cara a futuras acciones de mejora.

***Palabras clave:*** ingeniería mecánica y diferencias de género, medición de actitudes, análisis factorial multi-grupo, teoría de respuesta a los ítems.

In recent decades, a considerable decrease has been observed in the demand for STEM (science, technology, engineering and mathematics) university degrees in developed countries in general (e.g., Byars-Winston & Canetto, 2011; Knight et al., 2012; NSRCG, 2013) and Spain in particular (CRUE, 2008; Engynicat, 2008). Furthermore, this trend seems to be specially pronounced in the case of engineering degrees, particularly mechanical engineering (NSRCG, 2013; Ferrando et al., 2012).

Apart from this general decrease, mechanical engineering (ME) suffers from a well-known specific problem: the lack of interest shown by female students (Byars-Winston & Canetto, 2011; Hill, Corbett, & Rose, 2010; Medina, 2004; Knight et al., 2012). In Spain, the importance of this specific problem is quite clear (e.g., CRUE, 2008). After a timid increase in the 1970s, the percentage of women studying ME in Spanish universities has stabilized around values below 10% (see Ferrando, 2012, for a summary).

The decrease in the demand of the degree together with the lack of interest shown by women is creating a serious problem. For example, it is not clear that the existing engineering workforce in the USA will be able to maintain the present level of competitiveness (NSRCG, 2013). The problem has been studied by a wide range of research projects, most of which have two clear purposes: (a) to assess the situation and identify the main sources of the problem, and (b) to propose improvements. At the international level, these projects have been initiated either by the government or private foundations (for example, the ADVANCE initiative). In Spain, most of them have been carried out by universities: for example, the "Enginycat" (2008) project in Catalonia, the "Valentina" program in Valencia (Díez, 2010), and the proposals by Medina (2004) in Jaen and del Río (2009) in Madrid. The reviewed proposals, however, do not focus specifically on ME but on general engineering or, even more broadly, on STEM degrees in general.

The present study is also part of a research project that has been carried out at the Rovira i Virgili University (URV) and which focuses specifically on ME. The purposes of the project are: (a) to make a diagnosis of the situation, and (b) use the results to make

Received: January 11, 2013 • Accepted: May 17, 2013
Corresponding author: Pere Joan Ferrando Piera
Facultad de Ciencias de la Educación y Psicología
Universidad Rovira i Virgili
43007 Tarragona (Spain)
e-mail: perejoan.ferrando@urv.cat

a series of proposals for high schools designed to increase the demand of the degree, particularly among women. In this paper we shall discuss the first purpose, and the diagnosis will be made using a psychometric approach.

The initial predictions of the study were made on the basis of (a) the experience of the ME teachers at the URV and (b) the results of a series of semi-structured interviews administered to women who were studying the ME degree or who were professional engineers (Ferrando et al., 2012). The general starting point was that the low demand was partly because of an incomplete or distorted perception of ME as a career. This point can be divided into two more specific assumptions. First, high school students' knowledge about the competences and tasks of ME is at least incomplete and perhaps largely incorrect. Second, these students' perception of the work environments of ME professionals is negatively distorted, and this distortion is expected to be more pronounced in women. As far as the first assumption is concerned, we note that in Spain, the image of ME is closely related to the world of cars and motorcycles, particularly competition. In fact, however, only a small percentage of mechanical engineers work in this environment. And with regards to the second point, we believe that the work environments of ME are perceived as dirty, dangerous, physically demanding, and possibly hostile (see also Medina, 2004). Again, however, most mechanical engineers work in offices.

The psychometric assessment of the two above-mentioned specific points prompted us to develop two scales that we shall call "Tasks" (T) and "Perceptions" (P). Items on the T scale refer to tasks or activities that may or may not be specific to ME, and the respondent must decide whether or not they are. Items on the P scale are more general, and refer to the work environment and the personal characteristics of ME professionals. As well as the degree of specificity, the scales also differ in their degree of concreteness-abstraction (the P items are more abstract) and in the denotative-connotative dimension (Osgood, Suci, & Tannembaum, 1957), because P items reflect respondents' mental representation of ME as a career more than T items. The main difference, however, is that the T scale is mainly a maximum-performance objective measure (Cronbach, 1990) because its responses are naturally scored as correct-incorrect (competences in ME are defined by law). In contrast, the P scale is more similar to a typical-response measure (Cronbach, 1990) and can be considered to be an attitude scale. This last distinction, however, is not so clear-cut because the perceptions endorsed in the P scale can be realistic or distorted to some extent.

Two basic requirements are needed if T and P measures are to be used to assess (a) knowledge and perception at the individual and group level, and (b) potential gender-related differences in these two variables. First, they must both behave as essentially unidimensional measures and have some degree of measurement accuracy. Second, ideally their items should be gender invariant. Differential item functioning (DIF, see Muñiz, 1990) would mean that these items have different measurement properties in men and women, which would make gender-related comparisons more complex. Overall, then, the study had to be carried out in four stages: (a) development of the T and P measures, (b) assessment of their psychometric properties (unidimensionality, accuracy and invariance) in a representative sample, (c) comparison of the estimated levels in knowledge and perception in men and women, and (d) assessment of responses to individual items in order to identify the main sources of distortion and lack of knowledge.

To close this section, we shall summarize the main contributions of the present study. First, our study focuses specifically on ME. Second, most previous studies focused on external sources: for example, stereotypes, gender-related prejudices, differential work and family demands (Del Río, 2009; Hill, Corbett, & Rose, 2010; Medina, 2010; Spencer, Steele, & Quinn, 1999). In contrast, our starting points focus on knowledge and perception of ME as a career. So, we believe that the present results can complement those obtained in previous studies. Finally, our study uses a rigorous psychometric methodology and provides measurement instruments that can be used in future studies in Spain.

## Method

### *Participants and procedure*

The sample was collected from ten high schools in the province of Tarragona, and can be considered to be representative of the students that are admitted to the URV every year. The initial sample consisted of 496 high school students aged between 16 and 18, who were studying science or technology subjects. As explained below, inconsistent response patterns were removed from the data, so that the final sample consisted of 408 participants: 159 women and 249 men. It is noted that the percentage of respondents of each gender in the trimmed sample was similar to that in the original sample (about 39% women).

Questionnaires were administered in paper and pencil format in classroom groups and always by the same examiner. Administration was voluntary and anonymous: the only data requested were gender and age.

### *Instruments*

Items on the T (18 items) and P (9 items) scales were developed by one of the authors on the basis of (a) his experience as head of the ME degree, (b) the results of preliminary assessments to students, and (c) the semi-structured interviews referred to above. The content validity of the T items can partly be objectively assessed because, as mentioned above, the competences of a mechanical engineer are defined by law. The content validity of the T and P items was also assessed using inter-judge agreement. A committee made up of three experts (all of whom were mechanical engineers) assessed the initial pool of items, and the chosen items were those in which complete agreement was reached. The content of these items can be seen in Table 2.

In both scales, the response format is YES-NO. One item in each scale (T1 and P7) was used solely to assess response consistency and was not included in further analysis. On the basis of points (a) and (b) above, the content of these 'validity' items was considered too obvious to a high school student studying scientific or technical subjects. So, erroneous or distorted responses to the 'validity' items were considered to be indicators of response inconsistency, and the corresponding pattern was omitted from the data matrix.

### *Data analysis*

Given the different characteristics of the T and P scales, they were analyzed separately, and the psychometric model for the analysis was determined on the basis of the response format (binary) and the purposes of the study. In both T and P, the chosen model was the multiple-group extension of the two-parameter normal-ogive model (Muthén & Christofferson, 1981). Even though this model is most commonly used in its item response theory parameterization formulation, we decided to use the factor-analytic formulation for three reasons. First, the model-data fit assessment of the factor model the unidimensionality assumption to be rigorously assessed. Second, the invariance in the item measurement properties (i.e., lack of DIF) is also better assessed by using this modeling. Finally, the model provides estimates of the structural parameters at the group level (i.e., group means and variances on the corresponding traits) that enable gender differences to be strictly assessed.

In the factor-analytic parameterization of the multiple-group model, between-group differences in thresholds/intercepts are considered to indicate uniform DIF, whereas differences in the loadings indicate non-uniform DIF (e.g., Hernández & González-Romá, 2003). In accordance with this criterion, we consider that between-group invariance in both thresholds and loadings is a sufficient condition for assuming that the measure under scrutiny is free from DIF (see Little, 1997). This condition is known as "strong invariance" in the factor-analytic literature (Millsap & Meredith, 2007).

In the binary case, fitting a multiple-group model with the strong invariance restrictions allows the relative means in each group to be estimated (in fact the model is over-determined for this purpose) and, therefore, the potential gender differences to be assessed. The most usual procedure consists of setting one of the means to zero (in our case that of women's group) and freely estimating the mean of the other group together with its corresponding standard error (Muthén & Christofferson, 1981).

The model described above was fitted separately to T and P data by using WLSM estimation as implemented in the Mplus version 5.1 Program (Muthén & Muthén, 2007). As well as the chi-squared statistic, the following goodness-of-fit indices were used: (a) root mean square error of approximation (RMSEA) with its 90% confidence interval (Browne & Cudeck, 1993), (b) comparative fit index (CFI), and (c) nonnormed fit index- Tucker-Lewis index (NNFI-TLI) (see Hu & Bentler, 1999).

## Results

Table 1 shows the model-data fit results for both measures as well as the mean estimates in each group. According to the literature (Browne & Cudeck, 1993; Hu & Bentler, 1999; MacCallum & Austin, 2000), the model-data fit can be regarded as excellent in the case of P and acceptable in the case of T. Overall, we consider that the strong invariance condition is met in both scales. So, we conclude that there are no items with gender-related differential functioning or, more conceptually, that men and women interpret the items in the same way, and that, as measures, the items behave essentially in the same way in both groups.

The structural results do not support our starting assumptions on gender-related differences. Confidence intervals around the means which are freely estimated clearly show that there are no significant differences in any case. So, the results suggest that both the knowledge of tasks and the perceptions of work environments are essentially the same in men and women.

For both T and P, Table 2 shows the item measurement parameter estimates: discriminations and difficulties. Discrimination parameters are the standardized weights or loadings ($\lambda$). As for difficulties, we decided to report the means or proportions because

*Table 1*
Multiple-group analysis: Goodness of fit of the strong invariance model and group structural parameter estimates (means)

| (a) T Scale | | | | | | |
|---|---|---|---|---|---|---|
| **Model** | **$\chi^2$** | **df** | **RMSEA** | **90% C.I.** | **CFI** | **NNFI** |
| Strong I | 492.51 | 254 | 0.065 | (0.056;0.074) | 0.92 | 0.92 |
| | | | **Women** | **Men** | | |
| θ mean and 90% C.I. | | | 0 (fixed) | 0.011 (-0.219;0.241) | | |
| **(b) P Scale** | | | | | | |
| **Model** | **$\chi^2$** | **df** | **RMSEA** | **90% C.I.** | **CFI** | **NNFI** |
| Strong I | 60.82 | 46 | 0.040 | (0.000;0.065) | 0.97 | 0.97 |
| | | | **Women** | **Men** | | |
| θ mean and 90% C.I. | | | 0 (fixed) | 0.159 (-0.121;0.439) | | |

in this case they provide more information than the corresponding thresholds. Finally, given that the condition of strong invariance is met and that there are no significant group differences, only a single pooled estimate of each parameter common to both groups is reported.

Before the factorial solutions are interpreted, it is noted that overall measurement accuracy is acceptable in both scales even though some loadings (particularly in T) are quite low. The reliability estimates of the factor scores were 0.83 (T) and 0.80 (P).

The T structure can only be interpreted if it is taken into account that all of the items have been scored in the same direction: higher scores mean a higher level of knowledge. So, the difficulty indices are the proportions of correct responses in the corresponding items. To start with, it seems clear that the structure is "positive manifold", as is to be expected if all items measure a general dimension of knowledge in this domain. Closer inspection, however, reveals interesting results. First, there is a cluster of items with both high loadings and high proportions of correct responses (3, 5, 13, 15, 16 and 17) in which the level of knowledge seems to be clear. The content of these items refers to manual tasks specific to the profession of mechanic and even blacksmith, which respondents consider to be related to ME (high loadings) but, at the same time, understand that they are not actually tasks of an engineer.

At the other extreme there is a cluster of items with very low loadings and high proportions of incorrect responses. In principle, low discriminations might indicate problems of understanding due to item complexity, ambiguity, abstractness or length (Ferrando & Demestre, 2008). Inspection of the stems, however, suggests that this is not the case. Rather, it seems more plausible to simply assume that respondents perceive these items as being far removed from the dimension that is measured. The most extreme stems in this cluster refer to: organizing special transports (Item 8), organizing the manufacturing process of computer keywords (Item 9) and, above all, designing the trial for an artificial knee (Item 11). In these items, the proportions of incorrect responses are around 60-70%. So, overall, these items are not only probably perceived as quite unrelated to a career in ME but many respondents also believe that the corresponding tasks do not belong to ME.

*Table 2*
Item parameter estimates based on the strong-invariance model

| Item | Content | $\lambda_{zc}$ | Pc |
|---|---|---|---|
| | **T Scale** | | |
| T1 | Diseñar un motor* [ Designing an engine] | – | – |
| T2 | Proyectar una lavadora [Designing a washing machine] | 0.20 | 0.62 |
| T3 | Cambiar el aceite de un automóvil [Changing the oil of a car] | 0.79 | 0.80 |
| T4 | Apretar tornillos de máquinas [Tightening screws of a machine] | 0.75 | 0.64 |
| T5 | Soldar tuberías [Welding pipes] | 0.64 | 0.80 |
| T6 | Calcular la estructura de una nave industrial [Calculating the structure of an industrial warehouse] | 0.30 | 0.74 |
| T7 | Proyectar ascensores [Designing elevators] | 0.15 | 0.86 |
| T8 | Organizar y dirigir transportes especiales [Organizing and directing special transport] | 0.07 | 0.35 |
| T9 | Organizar la fabricación de teclados de ordenador [Organizing the manufacture of computer keywords] | 0.13 | 0.39 |
| T10 | Proyectar instalaciones de aire acondicionado [Designing air-conditioning facilities] | 0.10 | 0.60 |
| T11 | Definir el ensayo de una prótesis de rodilla [Defining the trial of an artificial knee] | 0.02 | 0.27 |
| T12 | Revisar las máquinas de un gimnasio [Revising gym equipment] | 0.63 | 0.80 |
| T13 | Cambiar personalmente la pala de una excavadora [Personally changing a bulldozer blade] | 0.70 | 0.78 |
| T14 | Pulir elementos metálicos [Polishing metal components] | 0.68 | 0.72 |
| T15 | Desatascar tuberías de una vivienda [Unblocking pipes at home] | 0.73 | 0.93 |
| T16 | Golpear con un martillo [Using a hammer] | 0.84 | 0.90 |
| T17 | Engrasar engranajes [Greasing gears] | 0.84 | 0.70 |
| T18 | Arreglar el reloj de un campanario [Mending the clock in a clock tower] | 0.38 | 0.68 |
| | **P Scale** | | |
| **Item** | **Content** | $\lambda_{zc}$ | **Pc** |
| P1 | Hace falta tener un carácter especialmente fuerte para dirigir equipos de obreros poco cualificados [You need to have a specially strong character to direct teams of unskilled workers] | -0.14 | 0.59 |
| P2 | Se trabaja casi siempre en entornos con riesgo de accidentes físicos [You almost always have to work in physically dangerous environments] | -0.64 | 0.67 |
| P3 | Se puede uno dedicar al diseño de productos [You can dedicate your time to designing products] | 0.49 | 0.15 |
| P4 | Se trabaja casi siempre en entornos sucios [You almost always have to work in dirty environments] | -0.73 | 0.83 |
| P5 | Casi siempre hay que llevar como vestido un mono azul [You almost always have to wear blue overalls] | -0.84 | 0.83 |
| P6 | Se trabaja mucho con ordenadores[ [You do a lot of work on the computer] | 0.50 | 0.14 |
| P7 | Se utilizan modelos físico-matemáticos para desarrollar el trabajo * [You have to use physical and mathematical models to do the work*] | – | – |
| P8 | Hace falta una fortaleza física importante para desarrollar las funciones propias del oficio [You need to be very strong physically to work as a professional mechanical engineer] | -0.58 | 0.80 |
| P9 | El entorno de trabajo principal son despachos y oficinas [You work largely in offices] | 0.64 | 0.52 |

Note. $\lambda_{zc}$: pooled standardized loadings; *Pc*: pooled proportion of endorsement ; *: validity item

The P solution also agrees with the expectations and, furthermore, the distribution of the algebraic signs of the loadings clearly suggests a bipolar structure. In accordance with our initial views, we interpret a positive pole related to ‘clean’ environments (computer-assisted design, office work) and a negative pole related to potentially hostile, dangerous and ‘dirty’ environments. Furthermore, inspection of item discriminations and difficulties (proportions of item agreement) suggests that, in this group of respondents, the negative perception is considerably stronger than the positive perception. In particular, note that the strongest loading is on the idea of ‘blue overalls’. Also remarkable is the percentage of respondents who relate ME to physically dangerous (70%) and “dirty” (80-85%) environments.

## Discussion

This study provides results that, as intended, may give rise to actions of improvement. Before discussing them, however, we shall focus on some of the limitations to the generalizability of the study. First, the sample used here can be considered to be representative of the population of high school students in the province of Tarragona studying science and technology subjects. However, there is no guarantee that this sample is also representative of the entire Spanish population. Generalization of the results is a matter for further research and, in this regard, the measurement instrument developed in this study can be a useful tool.

The fact that the sample used was restricted to the students studying science and technology is also a limitation, because much richer results might have been obtained by considering and comparing all the high school specialties. In this regard, we note that the study arose out of the need to improve a situation. So, we focused only on the potential 'consumers' of our subject.

The results obtained here can be summarized in three main points. First, there is a cluster of tasks traditionally related to ME that respondents tend to identify quite clearly. However, in other not so obvious tasks, the general lack of knowledge is considerable. Second, the perception of work environments and characteristics related to a career in ME is generally negative, and appears to be associated with stereotypes that are no longer real. Finally, contrary to expectations, no noticeable gender-related differences appear to exist regarding tasks and perceptions.

The first two points above agree with the expectations of the study. However, the results obtained add new and useful information. Thus, although the knowledge of 'traditional' tasks is generally acceptable, we note, for example, that about 40% of respondents believe that designing a washing machine is not a task for a mechanical engineer. However, it clearly is: here, we are referring to an industrial machine. As for the "not so obvious" tasks, lack of knowledge is mostly associated with the most usual tasks of ME at present (transport and manufacturing) or the ones that are most active or that have the brightest future (biomechanics). As regards the second point, the view of a career in ME is negatively distorted, as expected. However, it is interesting to note that the negative views mostly focus on the ideas of danger, physical demands, "dirtiness" and, above all, the belief that ME is a blue-collar profession.

Even though respondents tend to show a considerable lack of knowledge and a negatively distorted view of the career in general, there appear to be no significant gender-related differences on these points. So, in principle, our results cannot explain (at least directly) the lack of interest women show for ME. One possible explanation, which focuses on the P results, is that the 'negative' view differentially affects men and women. In other words, both men and women tend to perceive ME as 'dirty', 'hard', 'dangerous' and 'blue-collar'. However, this perception does not particularly worry men, but women find it sufficiently important not to seriously consider ME as a career. Indeed, this warrants further research.

To sum up, we believe that the results discussed so far are directly linked to potential actions of improvement to be carried out in the second stage of the project. As far as the tasks of a mechanical engineer are concerned, if information were provided about computer-assisted design, organization, transport, and biomechanics, the interest shown by students in general and women in particular might increase. After all, these tasks are neither 'dirty' nor 'blue-collar'. And as far as the perceptions of mechanical engineering are concerned, action should be taken to break away from the negative stereotypes and to spread the real image of a modern-day career in mechanical engineering.

## Acknowledgements

This study has been partially supported by a fund from the "DASC - Departament d'Acció Social i Ciudadanía, Institut Català de les Dones de la Generalitat de Catalunya" within the program "Subvencions per a treballs de recerca en matèria d'estudis de gènere i de les dones." Ref U-52/10.

## References

ADVANCE: http://www.mtu.edu/advance/.

Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 445-455). Newbury Park, CA: Sage.

Byars-Winston, A., & Canetto, S. (2011). Preface: Accomplishments and challenges for a diversity of women in science, technology, engineering and mathematics education and occupations. *Journal of Women and Minorities in Science and Engineering, 17*, 1-3.

Cronbach, L.J. (1972). *Essentials of psychological testing*. New York: HarperCollins.

CRUE: Conferencia de rectores de las Universidades españolas (2008). *La Universidad española en cifras [Spanish universities in figures]*. Observatorio Universitario.

Del Río, M. (2009). *Logros de las mujeres en la arquitectura y la ingeniería [Womens' achievements in architecture and engineering]*. Foro UPM.

Díez, J.L. (2010). *Proyecto Valentina [The Valentina project]*. Presentation at the 13th SIOU Meeting. Universidad Politécnica de Valencia.

Enginycat (2008). http://www.enginycat.cat.

Ferrando, F., Paleo, P., de la Flor, S., Urbina, C., & Gutiérrez-Colón, M. (2012). *Estudio sobre la baja presencia de las mujeres en los estudios de ingeniería mecánica [Assessing the low percentage of women in engineering degrees]*. Presentation at the 19th National Congress of Mechanical Engineering. Universidad de Castellón.

Ferrando, P.J., & Demestre, J. (2008). Características de forma y contenido que predicen la capacidad discriminativa en ítems de personalidad: un análisis basado en la Teoría de Respuesta a los Ítems [Characteristics of content and form that predict discriminating power in personality items: An IRT-based study]. *Psicothema*, *20*, 851-856.

Hernández, A., & González-Romá, V. (2003). Evaluating the multiple-group mean and covariance structure analysis model for the detection of differential item functioning in polytomous ordered items. *Psicothema*, *15*, 322-327.

Hill, C., Corbett, C., & Rose, A., (2010). Why so Few? Women in Science, Technology, Engineering and Mathematics. AAUW: Washington. http://www.eric.ed.gov/PDFS/ED509653.pdf.

Hu, L.T., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1-55.

Knight, D., Lattuca, L.R., Yin, A., Kremer, G., York, T., & Kyoung R.H. (2012). An exploration of gender diversity in engineering programs: A curriculum and instruction-based perspective. *Journal of Women and Minorities in Science and Engineering, 18*, 55-58.

Little, T.D. (1997). Mean and Covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research*, *32*, 53-76.

MacCallum, R.C., & Austin, J.T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51*, 201-226.

Medina, M.A. (2004). *Mujeres... ¿Ingenieras?[Women... engineers?]*. Round table at the congress 25 Years of Gender Studies. Universidad Jaime I de Castellón.

Millsap, R.E., & Meredith, W. (2007). Factorial invariance: Historical perspectives and new problems. In R. Cudeck & R.C. MacCallum (Eds.), *Factor analysis at 100* (pp. 131-152). Mahwah: LEA.

Muñiz, J. (1990). *Teoría de respuesta a los ítems [Item response theory]*. Madrid: Pirámide.

Muthén, B., & Christoffersson, A. (1981). Simultaneous factor analysis of dichotomous variables in several groups. *Psychometrika*, *46*, 407-419.

Muthén, L.K., & Muthén, B. (2007). *Mplus user's guide*. *Fifth edition*. Los Angeles: Muthén & Muthén.

National Survey of Recent College Graduates (NSRCG) (2013). http://www.nsf.gov/statistics/recentgrads/. Last visited: 25th April 2013.

Osgood, Ch. E., Suci, G.J., & Tannenbaum, P.H. (1957). *The measurement of meaning*. University of Illinois Press.

Spencer, S., Steele, C., & Quinn, D. (1999). Stereotype threats women's math performance. *Journal of Experimental and Social Psychology, 35*, 4-28.