Rank ordering of questionnaire items using Wilks's statistic: an example of application to three populations

Sandrine Gaymard, Michel Maurin

CSCanada
Cross-Cultural Communication
Vol. 13, No. 1, 2017, pp. 1-9
DOI:10.3968/9138

ISSN 1712-8358[Print]
ISSN 1923-6700[Online]
www.cscanada.net
www.cscanada.org

Rank Ordering of Questionnaire Items Using Wilks' Statistics: An Example of Application to Three Populations

Sandrine Gaymard[a],*; Michel Maurin[b]

[a]University of Angers, Laboratoire de Psychologie des Pays de la Loire (LPPL) EA 4638, Angers, France.
[b]IFSTTAR-LTE (Formerly INRETS-LTE), Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux, (LTE, Laboratoire Transport Environnement), Bron, France.
*Corresponding author.

Received 10 November 2016; accepted 22 December 2016
Published online 26 January 2017

Abstract

Wilks' test (1938) constitutes a major contribution to statistics and in practical applications. It is admittedly a classical method but even so it is of real interest because it is highly convenient to use. We hope to illustrate this in the following article by means of a study in the field of traffic psychology. In this demonstration, we start from a "characterization questionnaire" filled in by three populations of drivers and we test the population effect working from categorical data and 3-way tables. The differentiating power of the items is then examined and ranked in decreasing value. Through this example, we thus show the scope of Wilks'statistics and their highly general import with qualitative or category-specific data, compared with other techniques.

Key words: Characterization questionnaire; Population comparisons; Qualitative or category-specific data; Wilks'statistics; Traffic psychology

Gaymard, S., & Maurin, M. (2017). Rank Ordering of Questionnaire Items Using Wilks' Statistics: An Example of Application to Three Populations. *Cross-Cultural Communication, 13*(1), 1-9. Available from: http//www.cscanada.net/index.php/ccc/article/view/9138
DOI: http://dx.doi.org/10.3968/9138

INTRODUCTION

Samuel S. Wilks, known as the *Statesman of Statistics* (Mosteller, 1964), largely contributed to the theory and practice of statistics (David & Morrison, 2006). In the year of his PhD (1932), Wilks proposed noteworthy extensions of Fisher, Hotelling, Neyman, Pearson and Wishart's concepts. Namely in the beginning, that is, during his "normal" period (from 1932 to 1937-38), he took up and significantly developed the likelihood ratio L (or l) that Neyman and Pearson introduced in 1930-31 (Wilks, 1932; Pearson & Wilks, 1933). Indeed, Wilks also made an essential contribution in 1938 by considering normal approximations for other multi-parametered statistical distributions and developing the asymptotical distribution of the transformed lambda distribution $W = -2\ln L$(or $-2\ln l$).

The other particularity of Wilks resides in his desire to go beyond pure mathematics. For his doctorate which contributed to the small sample theory, he minored in education. His works came to the attention of Professor Lindquist, Professor of Education at Iowa University, who had used the technique of matched groups in his experimental works on educational psychology (Lindquist, 1931). The variety of Wilks' works is the expression of his manifold interests in different fields. Notably, Stephan et al. wrote (1965, p. 943): "His wide range of activities were an expression of his deep interests in mathematics and education as well as national defense, social science and the integrity of research". Thus, Samuel S. Wilks widely contributed to the practical application of statistics and it was practical problems which influenced both his own publications and those of his students (Stephan et al., 1965).

In the field of study of social representations there exist several tools and methods to analyse and understand social thinking and practices (Moscovici, 1961/76; Gaymard & Tiplica, 2015; Gaymard, Tiplica, Koh, & Wong, 2016; Nzobounsana & Gaymard, 2010; Tsoukalas, 2006; Vergès, 2001; Wagner, 1996 ; Wagner et al., 1999). In the last few years, it has been possible to identify specific works in the field of traffic psychology dedicated to vulnerable users such as pedestrians. The development of environmental models of moving such as walking constitutes a problem for society with pedestrian safety at stake (International Transport Forum, 2012). The tools identified in this specific field concern for example, free associations (Gaymard, Boucher, Greffier, & Fournela, 2012), the Conditional Scripts Questionnaire (Gaymard & Tiplica, 2015), the characterization questionnaire (Gaymard, Andrés, & Nzobounsana, 2011), discourse analysis using filmed sequences (Gaymard, Boucher, Nzobounsana, Greffier, & Fournela, 2013) or the little stories (Gaymard, 2012). The advantage of these tools is that they enable the study of the interactions between drivers and pedestrians; these interactions are not taken into account in most simulation models (Tom, Auberlet, & Bremond, 2008). Working from a characterization questionnaire, Gaymard, Andrés and Nzobounsana (2011) compared the rank order of the themes among young French and Spanish drivers. The rank order of the themes shows that the Spanish seem more aware of pedestrian vulnerability. If, for the moment, the only publication taking into account the age variable and driving experience with this type of tool is a report (Boucher & Gaymard, 2009), studies based on oculometric data can be found (Borowsky, Shinar, & Oron-Gilad, 2010; Underwood, Phelps, Wright, Loon, & Galpin, 2005). For example Borowsky et al. (2010) examined the effects of age and driving experience on the capacity to detect risks while driving. The results show that experienced drivers are more apt to detect potential hazards (e.g., with pedestrians), which is not the case of young, inexperienced drivers.

A common point in studies of social representations, as in social sciences in general, is a substantial presence of qualitative or ordinal data. Wilks' statistics are very convenient for such data and surveys that use them; we consequently propose to deal with Wilks' criterion and statistics in this study. We are therefore interested here in the characterization of the pedestrian by comparing three populations of drivers differentiated by age, background and driving experience.

1. METHOD

1.1 Tool and Populations

1.1.1 Characterization Questionnaire

Using exploratory interviews, a characterization questionnaire was designed to study the representation that drivers have of pedestrians. The characterization questionnaire (Gaymard, Andrés, & Nzobounsana, 2011) is composed of 20 items which deal with different types of users (e.g., children, elderly) but which also describe their attitudes or behaviour (e.g., respectful, cross anywhere) (see Table 1).

Flament, a psychologist and methodologist known for his works on the analysis of similarity or Boolean algebra (1962, 1976), reintroduced the characterization questionnaire in order to verify the hypothesis of the existence of collective rank ordering of themes (or items) in a survey. This questionnaire is comprised of a list of items the number of which has to permit hierarchical block ranking. Thus, working from a questionnaire comprised of 20 items, the subjects are asked first of all to choose the 4 items most characteristic of the object studied (here the pedestrian). This choice is mandatory: Four items are necessary and not three or five. Then, he or she is asked to choose the 4 items least characteristic of the object studied from the remaining items. After this, the subject chooses the 4 items which are still slightly characteristic and finally the 4 which are slightly less characteristic. Each item is then coded from 1 to 5:5 if chosen as characteristic, 1 if chosen as non characteristic, and 3 if not chosen by the subject questioned. This questionnaire is thus of a "rectangular Q-sort" type, which differs from the classical Q-sort that seeks to approach a Gaussian Law. It is not the items with a distribution whose mode is in the central class which concern us but on the contrary those with a very disymmetrical distribution which gives priority to the dimension of "characteristic" or "non characteristic" (Vergès, 2001).

Table 1
List of the Items of the Characterization Questionnaire

Careful	Teenagers
Don't respect the green light (for the drivers)	Space sharing
Have time	Pedestrian crossing
Unpredictible	Elderly people
Fragile	Have priority
Children	Impose themselves
Cross anywhere	Courteous
Wait	Cross slowly on purpose
In a hurry	Respectful
Dangerous	Group

1.1.2 The 3 Populations

The population of the study consisted of two groups of students from lay (state-subsidised) and catholic (private) universities: 20 lay students (population A "Young A"), 21 catholic students (population B "Young B"), and one older group, 20 experienced drivers. University fees differ considerably between the state and the catholic universities, which explains the differences in background between the lay students and those from catholic universities, who are more privileged. Moreover the experienced drivers tested came from a similar background to the lay students. Population A, had an average age of 20.8 years (SD = 1.51). Population B, had an average age of 22.8 years (SD = 0.75). Population C had an average age of 44.75 years (SD = 4.37).

1.1.3 The General Methodological Framework

There are several ways to analyze qualitative or category-specific data. For instance, factorial analysis describes the respective locations of the variables, their discrete values and the individuals, and explains the relationships between variables, the interactions and/or structures (Doise, Clemence, & Lorenzi-Cioldi, 1992; Gaymard, 2012). There is also, more or less implicitly, a lot of scaling or scoring processing that provides numerical values for the different qualities (Sheppard, 1962; Carroll & Arabie, 1980; Masson, 1980), although some of it is '*a bit hair raising*' (Luce & Galanter, 1963) and is also meaningless in the sense of measurement theory (Suppes & Zinnes, 1963; Roberts, 1979). In some cases it is possible to develop measurement representations (Michell, 1990; Maurin, 2011; Maurin & Wang, 2011) but each one applies to a particular type of recording without any overall implementation. Data and multifactorial analyses use matricial linear algebra and many clustering techniques based upon numerous metrics or semi and/or ultra metrics between individuals and variables (Lebart et al., 1984), and frequently employ techniques close to functional equations (Aczél, 1966).

But fortunately Wilks' statistics and tests are convenient for qualitative data (Maurin, 2004). So in this study we avoid factorial analysis assumptions as well as more complicated measurement; thus we choose to make use of Wilks' statistics for the 3-way contingency tables and related multinomial distributions, at the same time using a more traditional statistical methodology.

1.2 Statistics

1.2.1 Wilks' Statistics and Test

In 1938, Wilks introduced very powerful statistics (or criterion) by considering the log likelihood ratio for statistical distributions $D(X, q)$ depending on a multidimensional parameter $\theta \in \Omega$. In the ratio l or $L = L0\max/L\max$ the denominator is the maximum likelihood for the full model, the numerator is the maximum likelihood under a null hypothesis $H0$ corresponding to a regular sub-space $\omega \subset \Omega$, and Wilks showed that under $H0$, $W = -2\ln L$ follows asymptotically a Chi-squared distribution with dim(W)-dim(w)df. Later the result was complemented by Wald in 1943 and Chernoff in 1954 (Cox & Hinckley, 1974; Kullback, 1967; Lehmann, 1959). Then any analytically regular $H0$ hypothesis about the parameter of a very large number of distributions can be tested, and this is notably the case for multinomial distributions.

1.2.2 The Use of Wilks' Statistics for the Data

The data constitute a 3-way table with a row factor $i = 1\ldots I$ for items, a column factor $j = 1\ldots J$ for response categories and a layer factor $k = 1\ldots K$ for populations, (I = 20, J = 5, K = 3), and they follow a multinomial distribution with $pijk$ parameters under the single condition $\sum_{ijk}\, p_{ijk} = 1$. Here we are mainly concerned with a difference (or not) between the populations as with differences (or not) between items and these interrogations may be quite resolved when considering some ad-hoc sub-spaces in the $pijk$ parameters' space.

- For the complete model $L = \prod_{ijk} pijk^{nijk}$,

$$\ln L = \sum_{ijk} nijk \ln pijk, \tag{1}$$

and the maximum likelihood estimators $p_{ijk} = nijk/n{+}{+}{+}$, $i = 1\ldots I$, $j = 1\ldots J$, $k = 1\ldots K$. Then we get classically

$$\ln L\max = \sum_{ijk} nijk \ln nijk - n{+}{+}{+} \ln n{+}{+}{+}, \tag{2}$$

(the notation + in the subscript means the summation of data for the corresponding subscript).

- For every $H0$ null hypothesis on parameters there is a $L0\max$ and its Wilks' statistic $W(3, \mathrm{H0}) = 2\,(\ln L\max - \ln L0\max)$.

The basic first one $H0$ is the mutual independence between factors with

$$pijk = pi{+}{+}\; p{+}j{+}\; p{+}{+}k,$$

$$L_0 = \prod_{ijk} (pi{+}{+}\; p{+}j{+}\; p{+}{+}k)^{nijk}$$

$$= \prod i\, pi{+}{+}^{ni++} \prod j\, p{+}j{+}^{n+j+} \prod k\, p{+}{+}k^{n++k}, \tag{3.a}$$

$p_{ijk} = ni{+}{+}/n{+}{+}{+}$, …. and

$$\ln L0\max = \sum i\, ni{+}{+} \ln ni{+}{+} + \sum j\, n{+}j{+} \ln n{+}j + {+} \sum k\, n{+}{+}k \ln n{+}{+}k - 3\, n{+}{+}{+} \ln n{+}{+}{+}. \tag{3b}$$

Then

$$W(3, \mathrm{H0}) = 2\,\{\sum_{ijk} nijk \ln nijk - \sum i\, ni{+}{+} \ln ni{+}{+} - \sum j\, n{+}j{+} \ln n{+}j{+} - \sum k^{n}{+}{+}k \ln n{+}{+}k + 2\, n{+}{+}{+} \ln n{+}{+}{+}\}, \tag{3c}$$

follows asymptotically a chi-squared distribution with $IJK - 1 - (I{-}1 + J{-}1 + K{-}1) = IJK - I - J - K + 2$ df; we denote it the D null hypothesis with its $W(3, \mathrm{D})$ related statistic.

- Besides, each population has its own mode of responses in a 2-way table crossing items and categories, and we observe a population effect when there is a statistical dependence between the joint item-category distribution and the population distribution. With this H0 joint independence $pijk = pij{+}\; p{+}{+}k$ we get

$$L0 = \prod_{ijk} (pij{+}p{+}{+}k)^{nijk} = \prod ij\, pij{+}^{nij+} \prod k\, p{+}{+}k^{n++k}, \tag{4a}$$

$Pij{+} = nij{+}/n{+}{+}{+}$, $P_{++k} = n{+}{+}k/n{+}{+}{+}$ and

$$\ln L0\max = \sum ij\, nij{+} \ln nij{+} + \sum k\, n{+}{+}k \ln n{+}{+}k - 2\, n{+}{+}{+} \ln n{+}{+}{+}. \tag{4b}$$

Then

$$W(3, H0) = 2\,\{\sum_{ijk} nijk \ln nijk - \sum ij\, nij{+} \ln nij + - \sum k\, n{+}{+}k \ln n{+}{+}k + n{+}{+}{+} \ln n{+}{+}{+}\}, \tag{4c}$$

follows asymptotically a chi-squared law with $IJK - 1 - (IJ + K - 2) = (IJ-1)(K-1)$ df; we denote it the Ak null hypothesis with its $W(3, \mathrm{A}k)$ statistic.

- The joint independence of the i factor with $pijk = pi{+}{+}p{+}jk$ is noted Ai, and $W(3, \mathrm{A}i)$ follows asymptotically a Chi-squared law with $(I\text{-}1)\,(JK\text{-}1)$ df.

1.2.3 The Case of 2-Way Tables

Of course Wilks' tools are also useful for 2-way tables and notably for the marginal data of an initial 3-way table. In this case the independence between row and column factors is the sole H0; for instance for the *ij* marginal distribution after summation on the *k* layer factor, then the Wilks' statistic

$$W(2, k+) = 2\{\sum_{ij} n_{ij+} \ln n_{ij+} - \sum_i n_{i++} \ln n_{i++} - \sum_j n_{+j+} \ln n_{+jk} + n_{+++} \ln n_{+++}\}, \quad (5)$$

follows a Chi-square law with $(I-1)(J-1)df$.

Next we concern ourselves with the differences between the populations of respondents (subscript *k*), and notably by means of items (subscript *i*). So we consider the $i = 1\ldots I$ sub-2-way tables with $m_{jk} = n_{ijk}$ data, $j = 1\ldots J$, $k = 1\ldots K$. In this case

$$\ln L\max = \sum_{jk} m_{jk} \ln m_{jk} - m_{++} \ln m_{++}, \quad (6)$$

and concerning the *j* and *k* factors' independence *H*0 we have

$$\ln L0\max = \sum_j m_{j+} \ln m_{j+} + \sum_k m_{+k} \ln m_{+k} - 2\, m_{++} \ln m_{++}. \quad (7)$$

Then we get other Wilks' statistics noted

$$W(2, i) = 2\{\sum_{jk} n_{ijk} \ln n_{ijk} - \sum_j n_{ij+} \ln n_{ij+} - \sum_k n_{i+k} \ln n_{i+k} + n_{i++} \ln n_{i++}\}, \quad (8)$$

with $(J-1)(K-1)$ *df*, and we check

$$\sum_i W(2, i) = 2\{\sum_{ijk} n_{ijk} \ln n_{ijk} - \sum_{ij} n_{ij+} \ln n_{ij} + - \sum_{ik} n_{i+k} \ln n_{i+k} + \sum_i n_{i++} \ln n_{i++}\} = W(3, Bi). \quad (9)$$

2. ILLUSTRATION

As our data are in a 3-way contingency table n_{ijk} with the items factor *i* (rows), the reply categories factor *j* (columns) and the drivers' populations factor *k* (layers), they are in accordance with the foregoing statistics. So they are an occasion to illustrate and implement them, and we may immediately test some hypotheses related to our topics.

2.1 The Main Effects

To begin, as a preamble, the basic mutual independence hypothesis *D* yields $W(3, D) = 669.4$, $df = 274$ and the usual normal approximation for a Chi-squared = 13.2 ; then the null hypothesis of mutual independence is rejected and we may conclude that there is a global effect between the factors.

In order to be more precise we go into the question of an effect between the populations which concerns the joint independence between the layer factor of populations and the couple of row and column factors (respectively item and reply categories), say the *Ak* hypothesis. This is one of the central points of our surveys. Here we get $W(3, Ak) = 192.3$, $df = 198$ and the normal approximation = - 0.26. This supposes that the null hypothesis of joint independence is accepted, and we have to conclude that there is not a population effect on the data. In the same way we get $W(3, Ck) = 196.7$, $df = 200$ and the normal approximation = - 0.14 ; then the null hypothesis *Ck* of the non-participation of the layer factor *k* in the data is also accepted.

We consider also the A*i* null hypothesis related to the presence of an item effect with the joint independence between item factor *i* and the two other factors *j* for categories and *k* for populations. Now we get $W(3, Ai) = 669.3$, $df = 266$ and the normal approximation = 13.5. Then inversely this joint independence is rejected and we observe an item effect in the data. In complement we can also observe this effect on the *ij*-factor marginal distribution, and test it with the related $W(2, k+)$ statistic. Here $W(2, k+) = 477.1$, $df = 76$ and the normal approximation = 18.6; then the null hypothesis of marginal independence between items and categories is rejected and this confirms the presence of an item effect.

Lastly the test of the *Bi* null hypothesis is in accordance with the previous tests. Here $W(3, Bi) = 192.3$, $df = 160$ and the normal approximation = 1.75; consequently the conditional hypothesis of independence between *j* and *k* factors given item factor *i* is accepted, ($1.75 < 1.96$ for a first order risk $a = 0.05$). This agrees with the non-effect of the population factor *k* (see $W(3, Ak)$ and $W(3, Ck)$ results) whatever the effect of the item *i* factor (cf. $W(3, Ai)$ and $W(2, k+)$ results).

Then, concerning the main effects, a global effect is observed but the populations are not differentiated in the answers of all the participants in the items as a whole. However, since the interest of the characterization questionnaire is to reveal a collective ranking of the items, the presence of an item effect meets our expectations more. Subsequently, it seems relevant to test the differentiating power of the items and to develop reflection on their degree of characterization.

2.2 The Differentiating Power of the Items and Their Degree of Characterization

There is no real population effect but since the items express a strong difference in the answer, we are led to question whether some of them do not distinguish the populations more than others. To do this, we introduce the differentiating power of an item *i* concerning populations *k*. In this case, the sub-2-way tables of the n_{ijk} are considered for each item *i* and the $W(2, i)$ statistics which serve to test the hypothesis of independence between the populations and the reply categories for each *i* are introduced. *W* statistics asymptotically follow the chi-squared law with 8*df*, but considering an analytical level and not a statistical one, this numerical value also indicates the departure between the item *I* answers. The higher the $W(2, i)$ the greater the item effect on the answers. It is therefore an indicator of the differentiating power of the item concerning population and with these powers the differences between items can be observed and ranked by order of power. Table 2 presents this differentiating power of the items in decreasing order.

The first is statistically significant, that is to say the independence $H0$ is rejected for a first order risk $a = 0.05$ but again what is important here is the numerical data and their order. Some items are seen to be greater, and the first five for example can be considered to be dominant items in terms of differentiation of the three populations.

Table 2
Items Ranked by Decreasing Differentiating Power

Item *i*	*W*(2, *i*)
Children	20.29
Dangerous	15.72
Impose themselves	15.66
Pedestrian crossing	14.93
Have priority	13.2
Have time	12.62
Unpredictible	10.35
Courteous	10.16
Wait	10.13
Space sharing	9.99
Elderly people	9.48
Fragile	8.67
Teenagers	7.07
Don't respect the green light	6.78
In a hurry	6.03
Careful	5.99
Group	5.35
Cross anywhere	4.31
Respectful	3.88
Cross slowly on purpose	1.75

These results are unexpected since, in studies on conditionality, "children" has always appeared as a consensual element within the populations: "A child must be respected" (Gaymard, 2007). Moreover, previous results have shown that the children did not belong to the most characteristic elements of pedestrian representation (Boucher & Gaymard, 2009; Gaymard, Andrés, & Nzobounsana, 2011). In this context, the problem is that the child and the elderly person belong to the most vulnerable road users but they do not appear as central elements of the representation (Gaymard, Boucher, Greffier, & Fournela, 2012). Concerning the other dominant items, the item "dangerous" had already been identified as more characteristic among young people (Boucher & Gaymard, 2009; Gaymard, Boucher, Greffier, & Fournela, 2012) and more characteristic of young French people (Gaymard, Andrés, & Nzobounsana, 2011); it may be thought that the differentiating power observed here is oriented in the same direction. In the previous studies, the item "impose themselves" although not among the most characteristic items of the pedestrian, is nevertheless more characteristic for young drivers. Inversely, in the earlier studies, the item "pedestrian crossing" is among the characteristic items of the pedestrian and the one most often quoted by the young drivers and the experienced ones (Gaymard, Boucher, Greffier, & Fournela, 2012). At the other end of this list, the item "cross anywhere" for example was one of the most characteristic of the pedestrian and does not differentiate the populations here; the item "crosses slowly on purpose" which is the least differentiating item, appeared in the previous works among the least characteristic items of the pedestrian.

At this stage based on the early results, it can not be concluded that there exists a link between the differentiating power of certain items and the fact that they are more or less characteristic of the pedestrian.

There is here an interesting corollary because we may extract a partial 3-way contingency table from original data when retaining the first 12 items from 'children' to "fragile". In doing so we build a posteriori a new table which shows a population effect when considering its own $W(3, Ak)$ statistic; here $W(3, Ak) = 151.2$, $df = 118$ and the normal approximation = 2.06, (2.06 > 1.96 for $a = 0.05$). Of course every sub-table with ordered items 1 to n, $n \leq 12$ also shows a population effect.

2.3 Two by Two Population Comparisons

Table 2 comes from the study of the differentiating power of the items for all the data of the three populations and the dominant items are those from the association of these three populations. It may also be done with only two populations; the initial 3-way tables have two layers $K = 2$. However it is useful to complete the notation with a subscript K and to note $WK(2, i)$ for $K = 2$ or 3; the $W3(2, i)$ df is 8 and the one for $W2(2, i)$ is 4.

Below, the power of differentiation of the items $W2(2, i)$ is calculated for pairs of populations 2 by 2. Table 3 classes the items by decreasing power $W2(2, i)$ for the 4 pairs of populations : young A and young B, young A and experienced, young B and experienced, and lastly all the young A and B and the experienced.

Again some of them are significant ($H0$ rejected for $a = 0.05$), but the main importance is the numerical indications and the presence of items in the top levels, and possibly some repeated presences.

When the populations are compared 2 by 2, and if the first 3 most differentiating items are taken, a finer reading of the differences between the populations is obtained. Thus it can be noted that the item "children", which is the most differentiating between the young lay students ("young A") and the group of young catholics ("young B"), is also the most differentiating between the experienced drivers and the young catholics. It can therefore be supposed that this differentiating power is more marked among the young catholics ("young B"). Concerning the items "dangerous" and "impose themselves", these results could corroborate previous studies which identified these items as being more characteristic of young lay drivers ("young A") (Boucher & Gaymard, 2009). However, at this stage, it is not possible to say more explicitly which direction the differences take.

Table 3
The Respective Orders of Items for Decreasing Differentiating Power Between Populations

Young A & B		YA & exper		YB & exper		YA+B & exper	
Item *i*	***W*2(2, *i*)**	**Item *i***	***W*2(2, *i*)**	**Item *i***	***W*2(2, *i*)**	**Item *i***	***W*2(2, *i*)**
Children	12.51	Dangerous	11.53	Children	15.14	Impose themselves	12.02
Pedestrian crossing	11.12	Impose themselves	8.59	Impose themselves	11.95	Dangerous	10.78
Teenagers	6.95	Elderly people	6.56	Have priority	11.29	Have priority	9.94
Fragile	6.00	Unpredictible	5.89	Dangerous	8.42	Children	7.78
Have time	5.88	Careful	5.60	Have time	8.41	Have time	6.74
Unpredictible	5.70	Wait	5.08	Space sharing	7.19	Elderly people	6.53
Wait	5.18	Pedestrian crossing	4.99	Pedestrian crossing	6.35	Space sharing	5.55
Courteous	5.06	Have priority	4.79	Courteous	6.23	Courteous	5.10
Dangerous	4.94	Have time	4.60	Wait	6.12	Wait	4.95
Space sharing	4.44	Space sharing	4.35	Elderly people	5.12	Careful	4.81
In a hurry	4.13	Don't respect the green light	4.20	Fragile	3.79	Unpredictible	4.65
Impose themselves	3.64	Courteous	4.03	Don't respect the green light	3.76	Don't respect the green light	4.30
Respectful	3.29	Children	3.53	Unpredictible	3.54	Pedestrian crossing	3.81
Have priority	3.26	Fragile	3.46	Group	3.42	Group	3.29
Elderly people	2.95	Respectful	2.54	In a hurry	3.07	Fragile	2.68
Cross anywhere	2.56	In a hurry	2.13	Careful	2.37	In a hurry	1.89
Don't respect the green light	2.48	Teenagers	2.12	Cross anywhere	2.03	Cross anywhere	1.75
Group	2.06	Group	2.03	Teenagers	1.74	Cross slowly on purpose	0.85
Careful	1.18	Cross anywhere	1.34	Cross slowly on purpose	0.68	Respectful	0.60
Cross slowly on purpose	0.90	Cross slowly on purpose	1.01	Respectful	0.28	Teenagers	0.12

DISCUSSION

In this study, the aim was to show the interest of using Wilks' test (1938). This is a traditional statistical method but it has the advantage of being of general import and having varied practical applications. During his career, Wilks effectively focused on the relationship between theoretical and applied statistics. He worked with the College Entrance Examination Board and the Educational Testing Service developing the standardized tests that have had a profound effect on American education. He also worked with Walter Shewhart on statistical applications in quality control in manufacturing.

In this article the aim was to detail the procedure of Wilks' statistics through a concrete example of application to the field of traffic psychology, which is concerned with the relationship between psychological processes and the behavior of road users. We started from a characterization questionnaire, which is a frequently-used tool in the field of social representation; it makes it possible to characterize an object of representation through collective ranking by means of block choices (Gaymard, 2003; Gaymard, Andrés & Nzobounsana, 2011; Vergés, 2001). Until now, Wilks' statistics have not been applied to this tool.

The object in study of the characterization questionnaire used in this article bears on the representation of the pedestrian by drivers, which is a real social problem given the development of "eco-friendly" means of transport. For example, numerous articles deal with the problem of pedestrian injuries and fatalities, which more particularly concern children and the elderly (e.g. Chong, Poulos, Olivier, Watson, & Grzebieta, 2010; Damsere-Derry, Ebel, Mock, Afukaar, & Donkor, 2010; Gaymard, Boucher, Nzobounsana, Greffier, & Fournela, 2013).

We gathered data from three groups of drivers: Young drivers from a lay university (A), young drivers from a catholic university (B) and experienced drivers. The answers of these three groups serve as a basis for our demonstration. In order to do this the Wilks' test is a major contribution in statistics and some points may be mentioned.

- First there are some aspects in relation to multinomial distributions and the loglinear model. The analysis of *n*-way contingency tables was initiated by Roy in 1956 (Agresti & Gottard, 2005; Kendall & Stuart, 1961), presented by Birch (1963) for instance, and became the loglinear model (Agresti, 1990; Goodman, 1970; Lindeman, Merenda, & Gold, 1980). The statisticians' attention was mostly concerned with the breakdown of *pijk*... parameters, or ln *pijk*..., in connexion with main effects, two-factor interactions…, as in ANOVA. They used the maximum likelihood estimators and their fine properties for the *pijk* but it does not seem that they have widely and clearly used Wilks' test, (however note the *G*2 (= *W*) statistic in Agresti explicitly referenced to Wilks). Moreover although some expressions such as *nij* ln *nij* or *pij* ln *pij* appear (Agresti, 1990; Kupperman, 1959; Nelder & Wedderburn, 1972), and *nijk* ln*nijk* (Birch, 1963, Kullback, 1967), their properties apparently have not been systematically employed in loglinear models for the tests of null hypothesis.

-Another aspect is related to some inflexions or shifts in statistical terminology. One of them concerns the criterion; indeed there is a "basic" and unavoidable discrepancy between data and models in Statistics, and this has been emphasized with the special new term of deviance, "*the quantity -2 L*max *which we propose to call the deviance*" (Nelder & Wedderburn, 1972, note their notation *L for log-likelihood*). It then occurs as a measure of this discrepancy and finally '*the deviance or deviance difference is just a log likelihood ratio statistic*' (McCullagh & Nelder, 1989). Despite their deviance in 1989 in what they called the deviance difference of 1972, this term became a new name for the L Wilks' statistic or criterion. The result is that Wilks' name largely disappeared from statistical texts.

Independently, for the maximum likelihood in *n*-way contingency tables there are some nice and meaningful analytical breakdowns (Agresti, 1990), and certain others because of the closeness with Shannon's entropy and its own properties (Kullback, 1967; Maurin, 2001). As a last argument, the numerical calculations for a 3-way table are very easy to use with a mere spreadsheet on today's personal computers, which is much more straightforward than the tables used in Kuppermans' time as mentioned in his talk (Kupperman, 1959).

This approach is the central technical point of our examination of data, and for all of the above reasons the *W* statistic is without any doubt an outstanding tool for 3 or *n*-way tables without making any scaled value assumptions on factor modalities. In parallel we also consider that the name of Wilks and the Wilks' test deserve to be acknowledged.

If there exists no study in the field of road safety which compares young drivers from lay and catholic universities, Gaymard (2006) showed the impact of this variable of membership in the representation of the elderly person. Thus in her study young students from a catholic university appeared "less charitable" towards elderly people than young students from a lay university. An interpretation proposed by the author to explain these differences was that lay students generally belong to the middle class or the underprivileged and many of them work in different contexts to help others. So lay students could have developed a greater spirit of solidarity than catholic students.

Wilks' statistics constitute an interesting method for the analysis of characterization questionnaires dealing with several populations since, from the differentiating power of the items, they enable reflection on their degree of characterization. Thus an item which is not characteristic with regard to other items when working with a single group can become so with regard to other groups. This aspect, which allows a group to be situated with regard to another group, is fundamental in social psychology and in the study of social representations. Wilks' statistics such as they are applied to the field of education, the army and manufacturing have a very wide field of applications has we have shown here by associating mathematics to the domain of traffic psychology.

REFERENCES

Aczél, J. (1966). *Lectures on functional equations and their applications*. New York/London: Academic Press.

Agresti, A. (1990). *An introduction to categorical data analysis*. New York: John Wiley & Sons.

Agresti, A., & Gottard, A. (2005). *Independence in multi-way contingency tables: S N Roy's breakthroughs and later developments*. Working Paper 2005/08. Università Degli Studi di Firenze.

Birch, M. W. (1963). Maximum likelihood in three-way contigency tables. *Journal of the Royal Statistical Society*, *Serie B*, *25*(1), 220-233.

Borowsky, A., Shinar, D., & Oron-Gilad, T. (2010). Age, skill, and hazard perception in driving. *Accident Analysis & Prevention, 42*, 1240-1249.

Boucher, V., & Gaymard, S. (2009). *Visibilité psycho-physique des piétons dans l'environnement routier* (p.118). Rapport Final du Projet Vipper.

Carroll, J., & Arabie, P. (1980). Multidimensional scaling, *Annual review of Psychology*, *31*, 607-649.

Chong, S., Poulos, R., Olivier, J., Watson, W. L., & Grzebieta, R. (2010). Relative injury severity among vulnerable non-motorised road users: Comparative analysis of injury arising from bicycle–motor vehicle and bicycle–pedestrian collisions. *Accident Analysis and Prevention*, *42*, 290-296.

Cox, D. R., & Hinckley, D. V. (1974). *Theoretical statistics*. London: Chapman & Hall.

Crundall, D., Underwood, G., & Chapman, P. (1999). Driving experience and the functional field of view. *Perception, 28*, 1075-1087.

Damsere-Derry, J., Ebel, B. E., Mock, C.N., Afukaar, F., & Donkor, P. (2010). Pedestrians' injury patterns in Ghana. *Accident Analysis and Prevention*, *42*, 1080-1088.

David, H. A., & Morrison, D. F. (2006). Samuel Stanley Wilks. *The American Statistician, 60*(1), 46-49.

Doise, W., Clemence, A., & Lorenzi-Cioldi, F. (1992). *Représentations sociales et analyses de données*. Grenoble: Presses Universitaires de Grenoble.

Evans, T., & Macdonald, W. (2002). *Novice driver situation awareness and hazard perception*. Retrieved from http://www.rsconference.com/pdf/rs020031.PDF?check=1

Flament, C. (1962). L'analyse de similitude. *Cahiers du Centre de Recherche Opérationnelle*, *4*, 63-97.

Flament, C. (1976). *L'analyse booléenne de questionnaire*. Paris, la Haye: Mouton.

Garay-Vega, L., Fisher, D. L., & Pollatsek, A. (2007). Hazard anticipation of novice and experienced drivers: Empirical evaluation on a driving simulator in daytime and night time conditions. *Transportation Research Record, 2009*, 1-7.

Gaymard, S. (2003). *La négociation interculturelle chez les filles franco-maghrébines: Une étude de représentation sociale*. Paris: L'Harmattan.

Gaymard, S. (2007). La représentation de la conduite chez des jeunes conducteurs: Une étude de la conditionnalité routière. *Recherche Transports Sécurité, 97*, 339-359. doi: 10.3166/rts.97.339-359

Gaymard, S. (2012). Pedestrian representation through the analysis of little stories. *Psychology of Language and Communication, 16*(3), 185-200. doi: 10.2478/v10057-012-0013-9

Gaymard, S., Andrés, S., & Nzobounsana, V. (2011). La representación social y las emociones de los conductores Franceses y Españoles acerca de los peatones [Social representation and emotions in French and Spanish drivers toward pedestrians]. In F. Expósito, M. C. Herrera, G. Buela-Casal, M. Novo, & F. Fariña (Eds.), *Psicología jurídica. Áreas de investigación* (pp.435-442). Santiago de Compostela, Spain: Consellería de Presidencia, Xustiza e Administracións Públicas.

Gaymard, S., Boucher, V., Greffier, F., & Fournela, F. (2012). La perception du piéton par le conducteur: Aspects représentationnels et visuels. In S. Gaymard & A. Egido (Dir.), *Sécurité et facteurs humains dans les moyens de transports: Une approche multidisciplinaire* (pp.237-252). Paris: L'Harmattan.

Gaymard, S., Boucher, V., Nzobounsana, V., Greffier, F., & Fournela, F. (2013). La perception des piétons par les conducteurs: Corrélations entre les données d'un œil électronique et le verbatim des conducteurs [Drivers' perception of pedestrians: Correlations between the data of an electronic eye and drivers'verbatim]. *Canadian Journal of Behavioural Science/Revue Canadienne des Sciences du Comportement, 45*(2), 124-137. doi: 10.1037/a0025789

Gaymard, S., & Tiplica, T. (2014). Conditional respect towards the pedestrian: Differences between men and women and risk modeling by the Bayesian approach. *Quality & Quantity*, *48*(1), 91-110.

Gaymard, S., & Tiplica, T. (2015). Conditionality and risk for the pedestrian: Modelling with the Bayesian Networks. *International Journal of Injury Control and Safety Promotion*, *22*(4), 340-351.

Gaymard, S., Tiplica, T., Koh, P. P., & Wong, Y. D. (2016). Social thinking and risk between road users: The case of the pedestrian in two cultural contexts. *Cross-Cultural Communication*, *10*(12), 1-10.

Goodman, L. A., (1970). The multivariate analysis of qualitative data: Interactions among multiple classifications. *Journal of the American Statistical Association, 65*(329), 226-256.

Holland, C., & Hill, R. (2007). The effect of age, gender and driver status on pedestrians' intentions to cross the road in risky situations. *Accident Analysis and Prevention*, *39*, 224-237.

International Transport Forum. (2012). *Pedestrian safety, urban space and health*. OECD Publishing. doi: 10.1787/9789282103654-en

Kendall, M. G., & Stuart, A. (1961). *The advanced theory of statistics Vol 2 inference and relationship*. London: Charles Griffin and Company Limited.

Kullback, S. (1967). *Information theory and statistics*. New York: Dover Publications.

Kupperman, M. (1959). A rapid significance test for contingency tables. *Biometrics*, *15*(4), 625-628.

Lebart, L., Morineau A., & Warwick, K. M. (1984). *Multivariate descriptive statistical analysis: Correspondence analysis and related techniques for large matrices*. New York: John Wiley and Sons.

Lehmann, E. L. (1959). *Testing statistical hypotheses*. New York: John Wiley and Sons, Inc.

Lindeman, R. H., Merenda, P. F., & Gold, R. Z. (1980). *Introduction to bivariate and multivariate analysis*. Scott, Foresman, Glenview, IL.

Lindquist, E. F. (1931). The significance of a difference between "matched" groups. *Journal of Educational Psychology*, *22*, 197-204.

Luce, R. D., & Galanter, E. (1963). Discrimination. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol.1, pp.191-244). New York and London: John Wiley and Sons, Inc.

McCullagh, P., & Nelder, J. A. (1989). *Generalised linear model*. London: Chapman & Hall.

Masson, M. (1980). *Méthodologies générales de traitement statistique de l'information de masse*. Paris: CEDIC/Fernand Nathan

Maurin, M. (2001). *An entropy decomposition related to law's mixture, Bayesian inference and maximum entropy methods in science and engineering* (Vol.568, pp.132-138). 20th annual workshop, Gif-sur-Yvette, France, 2000, AIP conference Proceedings.

Maurin, M. (2004). Des tests de comparaisons de dépendance dans des cubes de contingence. *Journées de Statistique, SFdS, 23*(XXXVI).

Maurin, M. (2011) Un mesurage pour les nombres recueillis sur [0, 1] avec un curseur, au cours d'une présentation par couple de stimuli. *Mathématiques et Sciences Humaines/Mathematics and Social Sciences, 196*(4), 41-61.

Maurin, M., & Wang, X. (2011). Constructing meaningful numerical response when evaluating (dis)comfort perception in an univariate case. *Theoretical Issues in Ergonomics Science*. doi:10.1080/1463922X.2011.611263

Michell, J. (1990). *An introduction to the logic of psychological measurement*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Moscovici, S. (1961/76). *La psychanalyse, son image et son public* (2nd ed.). Paris: Presses Universitaires de France.

Mosteller, F. (1964). Samuel S. Wilks: Statesman of statistics. *The American Statistician*, *18*(2), 11-17.

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society, Series A, 135*, 370-384.

Nzobounsana, V., & Gaymard, S. (2010). Les analyses canoniques simple et généralisée linéaires: Application à des données psychosociales. *Mathématiques et Sciences Humaines/Mathematics and Social Sciences*, *189*(1), 69-101. doi:10.4000/msh.11653

Pearson, E. S., & Wilks, S. S. (1933). Methods of statistical analysis appropriate for k samples of two variables. *Biometrika*, *25*(3/4), 353-378.

Roberts, F. S. (1979). *Measurement theory, with applications to decision making, utility and the social sciences*. Boston: Addison-Wesley.

Sheppard, R. N. (1962). Analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrika*, *27*, 125-140.

Suppes, P., & Zinnes J. L. (1963). Basic measurement theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol.1, pp.1-76). New York and London: John Wiley and Sons, Inc.

Stephan, F. F., Tukey, J. W., Mosteller, F., Mood, A. M., Hansen, M. H., Simon, L. E., & Dixon, W. J. (1965). Samuel S. Wilks. *Journal of the American Statistical Association*, *60*, 312, 939-966.

Tom, A., Auberlet, J. M., & Brémond, R. (2008). Approche psychologique de l'activité de traversée de piétons. Implications pour la simulation microscopique du trafic urbain. *Recherche Transports Sécurité, 101*, 265-279.

Tsoukalas, I. (2006). A method for studying social representations. *Quality and Quantity, 40*, 959-981.

Underwood, N., Ngai, A., & Underwood, J. (2013). Driving experience and situation awareness in hazard detection. *Safety Science, 56,* 29-35.

Underwood, G., Phelps, N., Wright, C., Loon, E. V., & Galpin A. (2005). Eye fixation scanpaths of younger and older drivers in a hazard perception task. *Ophthalmic Physiological Optic, 25*, 346-356.

Vergès, P. (2001). L'analyse des représentations sociales par questionnaires. *Revue Française de Sociologie, 42,* 537-561.

Wagner, W. (1996). Queries about social representation and construction. *Journal for the Theory of Social Behaviour, 26*, 95-120.

Wagner, W., Farr, R., Jovchelovitch, S., Lorenzi-Cioldi, F., Marková, I., Duveen, G., & Rose, D. (1999). *Theory and method of social representations* [online]. London: LSE Research Online. Retrieved from http://eprints.lse.ac.uk/2640

Wilks, S. S. (1932). Certain generalizations in the analysis of variance. *Biometrika*, *24*, 471-494.

Wilks, S. S. (1938). The large sample distribution of the likelihood ratio for testing composite hypothesis. *Annals of Mathematical Statistics, 9*(1), 60-62.