2015, Vol.20 No 1, pp. 1-15
http://www.unioviedo.es/reunido/index.php/Rema
Revistas de la Universidad de Oviedo Versión electrónica
REUNIDO

Comparación de la aplicación vía Internet y lápiz y papel de cuestionarios para valorar la ira al volante en una muestra española

A Comparison of Internet-Based and Paper-and-Pencil Questionnaires in Assessing Driving Anger in a Spanish Sample

David Herrero-Fernández
Universidad de Deusto

RESUMEN

El presente estudio examinó las características de la aplicación vía Internet y vía lápiz y papel de los cuestionarios Driving Anger Scale (DAS) y Driving Anger Expression Inventory (DAX). Con este objetivo se comparó la estructura factorial, propiedades psicométricas y equivalencia de puntuaciones en dos muestras diferentes de conductores españoles (Internet, n = 201 participantes, y lápiz y papel, n = 329 participantes). En ambas condiciones el análisis factorial confirmatorio mostró un buen ajuste de 3 factores para el DAS (Avance impedido por otros; Conducción temeraria y Hostilidad directa) y de 5 factores para el DAX (Expresión verbal de la ira, Expresión física de la ira, Expresión de la ira mediante el vehículo, Expresión desplazada de la ira, y Expresión Adaptativa / Constructiva de la ira). Todos los factores de ambos cuestionarios correlacionaron positivamente entre sí, excepto el factor correspondiente a la expresión adaptativa de la ira, que lo hizo negativamente con el resto. Igualmente se observó que el método de aplicación de los instrumentos (Internet o lápiz y papel) no tuvo ningún efecto significativo en las pude los factores de ambos cuestionarios.

*Palabras clave***:** Rasgo ira al volante; Expresión de la ira al volante; Valoración a través de Internet; Lápiz y papel.

ABSTRACT

This study examined the Internet and paper-and-pencil version of the Driving Anger Scale (DAS) and Driving Anger Expression Inventory (DAX). With this aim, factorial structure, psychometric properties and score equivalences of the questionnaires were analyzed with two different samples (Internet, n = 201, and paper-and-pencil, n = 329) of Spanish drivers. In both conditions, confirmatory factor analysis showed a good fit of 3 factors for the DAS (Impeded Progress by Others, Reckless Driving, and Direct Hostility), and of 5 factors for the DAX (Verbal Aggressive Expression, Personal Physical Aggressive Expression, Use the Vehicle to Express Anger, Displaced Aggression and Adaptative / Constructive Expression). All of the DAS and DAX scales correlated positively with each other, except the adaptative form of expressing anger, which correlated negatively. The way of application (Internet vs. paper-and-pencil) had not significant effect on the scores of the scales of the DAS and the DAX.

Keywords: Driving Anger Trait; Driving Anger Expression; Aggressive Driving; Internet-Based Assessment; Paper-and-Pencil.

Contacto:
David Herrero Fernández
DeustoPsych I+D+i en Psicología y Salud. Universidad de Deusto
Avda. De las Universidades, 24, 48007, Bilbao (ESPAÑA - SPAIN)
Email: d_herrero@psicologos.com

1.- Introduction

1.1.- Internet based assessment

Internet-based assessment is becoming more frequent each day. Technical development allow researchers to design sophisticated experiments to collect data in many different areas (e.g. perception, learning, leadership, etc.). This methodology has shown many advantages, such as the access to large samples, that allows a greater external validity (Buchanan, 2000; Buchanan & Smith, 1999; Musch, Reips, & Birnbaum, 2000; Pettit, 1999); the lower experimental costs (Buchan, DeAngelis, & Levinson, 2005; Buchanan & Smith, 1999; Musch, et al., 2000; Sampson, 2000); the possibility of providing tools around the clock, without any time limit (Buchan, et al., 2005; Musch, et al., 2000); the emphasis on voluntary participation, which usually increases respondents' motivation (Buchan, et al., 2005; Buchanan & Smith, 1999; Musch, et al., 2000); the data entry stage is eliminated (Buchan, et al., 2005; Cook, Heath, Thompson, & Thompson, 2001; Sampson, 2000); the increase of versatility in the design of the tasks (Buchanan & Smith, 1999); the study of crosscultural topics (Fang, Wen, & Prybutok, 2014; Suarez-Balcazar, Balcazar, & Taylor-Ritzler, 2009), and the possibility of decreasing the influence of demand characteristics, observer bias and response bias (Davis, 1999). Other important advantage is the different survey modes of data collection via Internet. The most used methods are email, online through the World Wide Web, mobile and SMS-based surveys (Alam, Khusro, Rauf, & Zaman, 2014). All of them have showed the commented advantages over other the traditional methods (Christie, Dagfinrud, Dale, Schulz, & Hagen, 2014; Hunter, Corcoran, Leeder, & Phelps, 2013; Sutherland, Amar, & Laughon, 2013).

But Internet assessment has also some limitations. One of the most important refers to the fact that psychometric properties of paper-and-pencil and Internet versions of a questionnaire might not be comparable (McKee & Levinson, 1990; Moreland, Zeidner, & Most, 1992; Leslie, 2006; Meade, Michels, & Lautenschlager, 2007). Thus, it is necessary to test measuring instruments in this environment, because their construct validity can be altered.

There are questionnaires of many psychological topics that have been adapted from paper-and-pencil version to an Internet version. Most of them have shown similar psychometric properties, like the Career Key Interest Inventory (Buchan, et al., 2005) the MMPI (Hays & McCallum, 2005), and the Toronto Alexithymia Scale (Bagby, Ayearst, Morariu, Watters, & Taylor, 2014). However, some others have proved that the scores are influenced by the method of application. This effect has been observed in topics like social desirability (Richman, Kiesler, Weisband, & Drasgow, 1999), physical and sexual attractiveness (Epstein, Klinkenberg, Wiley, & McKinley, 2001); mood regulation (Fouladi, McCarthy, & Moller, 2002); administration time (Vispoel, Boo, & Bleiler, 2001); students' rating of instruction (Chang, 2005); reading and vocabulary skills (Pomplun & Custer, 2005; Pomplun, Frey, & Becker, 2002); substance use (Wang, et al., 2005); and self-focused rumination (Davis, 1999).

1.2.- Why driving anger must be studied?

Road accidents cause many deaths among young people, and they are influenced by three main variables: Human factor, road status and vehicle. Human factor explains significantly more variance in the prediction of road accidents than the other two (Evans, 1991). According to several studies, one of the most important variables because of its implication in the accidents is anger (Dahlen & Ragan, 2004; Deffenbacher, Filetti, Richards, Lynch, & Oetting, 2003; Deffenbacher, Lynch, Filetti,

Dahlen, & Oetting, 2003). The reason is that the emotional arousal labeled as anger has a negative influence on some cognitive variables, like attention, perception and information processing, which influence the driver´s control of the vehicle while driving (Bone & Mowen, 2006; Pinto, 2001).

Driving anger has been considered as a personality trait (Deffenbacher, Richards, Filetti, & Lynch, 2005; Parker, Lajunen, & Summala, 2002; Underwood, Chapman, Wright, & Crundall, 1999), different but related to the general anger trait, to trait anxiety and to impulsiveness (Deffenbacher, Lynch, et al., 2003). The study of this trait is important because some researchers have shown that there is a relationship between driving anger and crash related conditions in simulator tasks (Deffenbacher, Lynch, Oetting, & Yingling, 2001).

Trait driving anger can be measured through the Driving Anger Scale (DAS) (Deffenbacher, Oetting, & Lynch, 1994), a questionnaire that identifies several situations that provoke anger in drivers. It consists of 33 items and has been adapted with samples from several countries. In addition, the DAS has a short version of 14 items which have been extracted from the large version because of their high factorial loads. It has been adapted specifically only with a Spanish sample (Herrero-Fernández, 2011a), showing that the Spanish drivers are angered by three general situations: Impeded Progress by Others, Reckless Driving and Direct Hostility.

However, trait anger assessment is necessary but insufficient, since two drivers with the same level of anger can express their road anger differently. One of the questionnaires that have been created to measure the anger expression is the Driving Anger Expression Inventory (DAX) (Deffenbacher, Lynch, Oetting, & Swaim, 2002). This questionnaire has been also adapted with a Spanish sample (Herrero-Fernández, 2011b). This research found a fit of five ways of expressing anger: Verbally, Physically, Using the Vehicle, Displacedly and Adaptatively. The only difference with the original version is that in the latter the "Displaced Expression" factor was suppressed because of its low reliability. Therefore, "driving anger" must be assessed the experienced anger as well as the anger expression.

The main goal of this research was to explore the potential of Internet-based research methods in the study of driving anger, in comparison with the paper-and-pencil method. With this aim, the factorial fit (confirmatory factor analysis) of the DAS and DAX through Internet and paper-and-pencil and the scores of the two samples will be compared. Thus, if the fit is equal in the two ways of application for each questionnaire, and there are not differences in the scores of the two samples, the driving anger assessment through Internet will be able to be considered to the clinical and research practice. Finally, the validity of the two versions of each questionnaire was analyzed. In this case, the validity was tested by correlating the DAS and DAX scores, because of the similarity of these constructs (Oren, Kennet-Cohen, Turvall, & Allalouf, 2014). This is one of the most used validity type according to the newest standards (to see a complete revision of the new standards, see: Lane, 2014; Padilla & Benitez, 2014; Rios & Wells, 2014; and Sireci & Faulknen-Bond, 2014).

2.- Method

2.1.- Participants

Two independent samples were studied in this research. The assignment of participants to the Internet or to the Paper-and-Pencil condition was made previously for convenience.

Three hundred and thirty eight questionnaires were delivered in paper and pencil way of application, with 329 of them (97.3%) being correctly completed. This condition was integrated by 66 males and 263 females (n = 329). Their age ranged from 18 to 57 (*Mdn* = 21.00).

Regarding the Internet application procedure, 312 questionnaires were sent by email to the participants, with 201 of them (64.42%) being correctly completed. This condition consisted of 105 males and 96 females (n = 201), with age ranging from 19 to 71 (*Mdn* = 27.00).

Finally, the only requirement to participate in the present research was to have diving license and to drive, at least, once a week.

2.2.- Instruments

Driving Anger Scale (DAS). The DAS (Deffenbacher, et al., 1994) is a five-point Likert scale (from 1 = Not at all to 5 = Very much) that assesses trait driving anger measuring the level of anger experienced when the driver is in the situation described by each item. There is a 14-item short form that has been adapted with a Spanish sample (Herrero-Fernández, 2011a), showing that it is divided in 3 factors: Impeded Progress by Others (α = .77), Reckless Driving (α = .66) and Direct Hostility (α = .87). These three factors can be summed into a global score (α = .84).

Driving Anger Expression Inventory (DAX). The DAX (Deffenbacher, et al., 2002) has been also adapted with a Spanish sample (Herrero-Fernández, 2011b). It is a four-point Likert scale (from 1 = Almost never to 4 = Almost always) of 50-item, which measures the way of expressing anger behind the wheel. The Spanish adaptation has evidenced five factors: Verbal Aggressive Expression (α = .91); Personal Physical Aggressive Expression (α = .79); Use of the Vehicle to Express Anger (α = .82); Displaced Aggression (α = .78), and Adaptative / Constructive Expression (α = .81). The desadaptative forms of expressing anger can be summed into the Total Aggressive Index (α = .92).

2.3.– Procedure

The order of the questionnaires was counterbalanced in each one of the two ways of application, so that approximately fifty percent completed firstly the DAS and then the DAX, and the other fifty percent did it in the opposite order. A single sheet with the two questionnaires was given to the paper-and-pencil group, while the Internet group received an email with the same format sheet attached. The Internet participants had been informed previously that they were going to receive the email with the questionnaires. Instructions were placed in the same questionnaires sheet, just before each questionnaire, both in the paper-and-pencil and the Internet conditions.

3.- Results

3.1.- Psychometric properties

The four models of questionnaire (DAS Internet, DAS paper-and-pencil, DAX Internet, and DAX paper-and-pencil) were analyzed through a confirmatory factor analysis. In order to guarantee the parsimony principle, two factorial structures for each questionnaire in each condition were tested, so that the simplest one would be kept if the two of them fitted equally well. The DAS was tested to compare its fit with one latent factor and with three latent factors: Impeded Progress by Others, Reckless Driving, and Direct Hostility, trying to replicate the results obtained in a UK sample

with the larger version, that fitted in three factors named similarly (Lajunen, Parker, & Stradling, 1998). The DAX was tested to verify its fit with two latent factors: Adaptative / Constructive Expression and Total Aggressive Expression, and with five latent factors: Verbal Aggressive Expression, Personal Physical Aggressive Expression, Use of Vehicle to Express Anger, Displaced Aggression and Adaptative / Constructive Expression (Deffenbacher, et al., 2002).

The goodness of fit analyzed indices for each model were the χ^2/df index, being acceptable a value lower than 5 (Wheaton, Muthén, Alwil, & Summers, 1977); the Root Mean Squared Error of Approximation (RMSEA), being reasonable any value equal or lower than .08 (Browne & Cudeck, 1993); and the Comparative Fix Index (CFI) and Non-Normed Fit Index (NNFI), being acceptable for these two values of .90 or more (Bentler & Bonnet, 1980). In addition to this, in order to test statistically the goodness of fit of the each pair of comparisons, the χ^2 test of contrast was used. The results of all these tests are shown in Table 1.

	χ^2/df	RMSEA	CFI	NNFI
DAS Int. [A]	3.70	.12	.87	.85
DAS Int. [B]	2.16	.07	.95	.94
$\lvert\chi^2_A - \chi^2_B\rvert$	123.62*			
DAS P-a-P [A]	5.42	.12	.86	.83
DAS P-a-P [B]	3.41	.08	.93	.91
$\lvert\chi^2_A - \chi^2_B\rvert$	185.30*			
DAX Int. [C]	4.34	.13	.78	.77
DAX Int. [D]	2.63	.08	.85	.84
$\lvert\chi^2_C - \chi^2_D\rvert$	2035.19*			
DAX P-a-P [C]	5.46	.12	.76	.84
DAX P-a-P [D]	3.01	.07	.90	.90
$\lvert\chi^2_C - \chi^2_D\rvert$	2893.56*			

INT: Internet. P-a-P: Paper-and-pencil.
A: One factorial structure. B: Three factorial structure. C: Two factorial structure. D: Five factorial structure.
** p < .001.*

Table 1. Confirmatory factor analysis for DAS and DAX

All the χ^2 tests were significant, indicating that the best fitted model was significantly different from the worst one. Therefore, the DAS fitted in three factors, both for the Internet and for the paper-and-pencil methods; and the DAX fitted in five factors, both for the Internet and for the paper-and-pencil methods. In the case of Internet condition the incremental indices were slightly lower than the cut-off. However, given the acceptable values for all the other indices, this model is permissible

(see Bollen, 1989; Raykov & Marcoulides, 2006). The descriptive statistics and internal consistence of each factor for each condition are shown in Table 2.

	INT n = 201			P-a-P n = 329		
	M	SD	α	M	SD	α
DAS						
Impeded Progress by Others	17.91	5.36	.80	19.01	4.89	.77
Reckless Driving	17.22	3.64	.70	17.36	3.42	.67
Direct Hostility	6.13	2.41	.88	6.70	2.28	.84
Total	41.27	9.62	.87	43.07	8.80	.85
DAX						
Verbal Aggressive Expression	23.56	8.20	.91	25.02	8.95	.92
Physical Aggressive Expression	11.30	2.64	.79	11.56	2.94	.81
Use of the Vehicle to Express Anger	14.71	4.31	.84	14.88	4.39	.84
Displaced Aggression	3.75	1.38	.77	4.21	1.71	.78
Adaptative/Constructive Expression	36.26	7.63	.83	37.22	7.00	.81
Total Aggressive Expression Index	53.32	13.00	.92	55.68	13.61	.92

Int.: Internet condition; P-a-P: Paper-and-pencil condition

Table 2. Descriptive Statistics and Internal Consistence (Cronbach´s Alpha) of the DAS and the DAX in both conditions.

In order to test the validity of the two questionnaires, the correlation coefficients between the DAS scales and the DAX scales were obtained, separately for each condition (Internet and paper-and-pencil). The results are presented in Table 3. Almost all the correlations were significant, and all of them were positive except the coefficients referred to Adaptative / Constructive Expression factor, which were negative and many of them statistically significant. This was coherent, because all of the factors refer to trait anger (DAS) and to desadaptative / negative forms of expressing anger (DAX), with the exception of Adaptative / Constructive Anger Expression factor.

	1	2	3	4	5	6	7	8	9	10
1. IPO		*.68****	*.45****	*.93****	*.29****	*.26****	*.32****	*.25****	*-.24***	*.37****
2. RD	.61***		*.39****	*.85****	*.24***	*.22***	*.23***	*.25***	*-.15**	*.30****
3. DH	.44***	.40***		*.65****	*.29****	*.21***	*.24***	*.26****	*-.21***	*.33****
4. DAS	.91***	.83***	.66***		*.32****	*.28****	*.32****	*.30****	*-.24****	*.40****
5. VAEX	.28***	.25***	.27***	.32***		*.54****	*.41****	*.19***	*-.16**	*.90****
6. PAEX	.33***	.20***	.20***	.31***	.55***		*.63****	*.20***	*-.23***	*.78****
7. UVEA	.36***	.16**	.18**	.31***	.46***	.71***		*.15**	*-.33****	*.73****
8. DIAG	.20***	.14**	.14**	.21***	.24***	.23***	.23***		*-.01*	*.32****
9. ACEX	-.22***	-.07	-.16**	-.22***	-.22***	-.31***	-.40***	-.12*		*-.26****
10. TAIN	.38***	.26***	.28***	.38***	.90***	.79***	.76***	.38***	-.34***	

IPO: Impeded Progress by Others; RD: Reckless Driving; DH: Direct Hostility; DAS: Total DAS Score; VAEX: Verbal Aggressive Expression; PAEX: Physical Aggressive Expression; UVEA: Use of the Vehicle to Express Anger; DIAG: Displaced Aggression; ACEX: Adaptative/Constructive Expression; TAIN: Total Aggressive Index.
The horizontal line between rows 4 and 5 split the table between DAS subcales and DAX subescales.
**p* < .05; ***p* < .01; ****p* < .001.

Table 3. Correlations between DAS and DAX scales, in both Internet (coefficients in italics) and Paper-and-Pencil ways of application.

3.2.- Differences by method of application

In order to examine potential differences between Internet and paper-and-pencil methods of application, a one way MANCOVA with method of application as factor was carried out for the DAS, and another one for the DAX. Age group (<30, 30-44 and >44) and Gender (males – females) were introduced as covariates. All the effect sizes (η^2) were interpreted according to Cohen´s criterion, wherein values between .01 and .04 correspond to a small effect size; between .05 and .14 correspond to a medium effect, and more than .14 correspond to a large effect (Cohen, 1988).

In the case of DAS, a multivariate effect for the method of application was observed, $F(3, 524) = 2.71$, $p = .044$, $\eta^2 = .01$, but there were no significant differences in the univariate analysis. The Age showed a significant effect, $F(3, 524) = 8.27$, $p < .001$, $\eta^2 = .05$. In the case of DAX, there was not multivariate effect for method, $F(5, 522) = 1.32$, $p = .253$, although there were multivariate significant effects for Age, $F(5, 522) = 7.42$, $p < .001$, $\eta^2 = .07$. The gender was not significant in any case. All the univariate analyses are shown in Table 4, with the Hochberg´s GT2 Post Hoc test. This test was used because the large differences in the sample sizes (Field, 2005). As it can be seen, young drivers scored higher than olders in all of the subscales of the DAS and DAX, except in the Adaptative / Constructive way of anger expression. Thus, in general drivers aged < 30 got higher scores than drivers aged 30 – 44 and > 44, while the only differences between these two last groups was in Verbal anger expression and in the Total desadaptative index of anger expression.

Factor	Age						*F*	η^2
	<30		31 – 44		>44			
	n = 386		n = 85		n = 59			
	M	SD	M	SD	M	SD		
DAS								
Impeded Progress by Others	19.04c	4.68	17.75	5.59	16.90a	6.40	6.00*	.02
Reckless Driving	17.66bc	3.19	16.61a	3.90	16.00a	4.41	7.92**	.03
Direct Hostility	6.84bc	2.25	5.91a	2.34	5.03a	2.27	19.51**	.07
Total	43.53bc	8.27	40.27a	9.93	37.93a	11.46	12.83**	.05
DAX								
Verbal Expression	25.78bc	8.91	22.82ac	7.38	18.24ab	5.39	22.80**	.08
Physical Expression	11.70c	3.13	10.95	1.67	10.68a	1.65	5.03*	.02
Use of the Vehicle	15.08c	4.55	14.80	3.98	13.12a	3.06	5.28*	.02
Displaced Expression	4.23bc	1.70	3.59a	1.09	3.44a	1.32	10.40**	.04
Adaptative Expression	36.85	7.01	35.85	7.70	38.31	7.98	2.01	.01
Total Desadaptative	56.78bc	14.46	52.16ac	11.13	45.47ab	8.62	20.04**	.07

Note: Differences by age, being <30 (a), 30-44 (b) and >44 (c), according to the Hochberg´s GT2 Post Hoc test.
*$p < .01$, **$p < .001$.

Table 4. Differences by age in DAS and DAX.

4.- Discussion

Nowadays it is common to adapt psychometrical scales from paper-and-pencil versions to Internet-based formats, since new technologies make easier the assessments of psychological topics. Their application needs to be adapted, since a paper-and-pencil based version may have different psychometric properties than the corresponding Internet version (McKee & Levinson, 1990; Moreland, et al., 1992). The aim in this research was to compare the two ways of assessing driving anger. In order to do this, the factorial fit of the DAS and DAX through Internet and paper-and-pencil was compared, as well as the scores of the two samples.

The present research has shown that confirmatory factor analysis of the short version of Driving Anger Scale (DAS) has replicated the results of the Spanish adaptation of this questionnaire (Herrero-Fernández, 2011a) both in Internet and in paper-and-pencil samples. Thus, it has been supported a factorial structure equivalent to that found with a UK sample (Lajunen, et al., 1998), composed by three factors: Impeded Progress by Others, Reckless Driving and Direct Hostility, as well for the

Internet sample as for paper-and-pencil sample. The original authors had presented this scale as a monofactorial short version (Deffenbacher, et al., 1994).

The Driving Anger Expression Inventory (DAX) fitted in a five-factor structure better than in a two-factor structure, both with Internet sample and a paper-and-pencil sample. However, there were two goodness of fit indices (CFI and NNFI) that scored underneath the established acceptance criterion of .90. The reason of this fact can be the small sample size for the DAX in the Internet condition, so it might be necessary to carry out other studies that confirm our results. Nevertheless, the main difference with the original research (Deffenbacher, et al., 2002) was that the Displaced Aggression scale was discarded in the original research because of its low reliability, whereas it has been accepted in our research because of its good result in the internal consistency. This effect was observed in the original Spanish adaptation of the DAX (Herrero-Fernández, 2011b).

The study of the concurrent validity has been carried out correlating all the scales of the DAS and of the DAX with each other, both for the Internet condition and for the paper-and-pencil condition. Results are concordant with the other studies (Dahlen & Ragan, 2004; Deffenbacher, Lynch, Deffenbacher, & Oetting, 2001; Deffenbacher, et al., 2002; Deffenbacher, White, & Lynch, 2004; Esiyok, Yasak, & Korkusuz, 2007), showing positive and significant coefficients between the DAS and the DAX scales. The only factor that correlated negatively with all the other scales (reaching statistical significance almost all coefficients) was the Adaptative / Constructive Expression scale. This fact seems consistent, because it is the only scale that indicates an adaptative way of behaving behind the wheel. In addition to this, the sign of coefficients are the same in all the cases, and the strength of each peer of correlations through the two conditions (Internet/paper-and-pencil) is very similar. These results confirm the validity for the two questionnaires for both application modalities.

After the factorial structure had been confirmed, the scores of the two samples were compared, being age and gender considered as covariates. The age groups were formed by aged <30, 30-44 and >44, following the criteria assumed in the works of Spanish adaptations (Herrero-Fernández, 2011a, 2011b). In the case of DAS, the results evidenced a multivariate difference for the method of application, but the effect size was very small and there was not any univariate difference. The age showed a significant multivariate difference, with a moderate effect size, and all the univariate comparisons were significant. Thus, in general the youngest drivers scored higher than older. These data replicate other studies, for example the Turkish adaptation of the DAX (Esiyok, et al., 2007), which showed that drivers aged less than thirty years scored higher than older drivers in Physical Expression, in Use of Vehicle to Express Anger and in Total Aggressive Index. In the Spanish adaptation of the DAX (Herrero-Fernández, 2011b), younger drivers scored higher than older drivers in each way of expression, with the exception of the Adaptative / Constructive Expression, with low to moderate effect sizes. Additionally, some studies have evidenced that aggressivity, as well as the risky behaviors behind the wheel, decrease with the age (Dahlen, Martin, Ragan, & Kuhlman, 2005; Elander, West, & French, 1993; Goehring, 2000; Parker, Reason, Manstead, & Stradling, 1995; Schwartz & Deffenbacher, 2002).

In conclusion, these results confirm that the method of application of the DAS and the DAX questionnaires is equivalent both in the Internet and in paper-and-pencil conditions, according to reliability and validity indices and to the magnitude of the scores, since there were no differences by the method of application. This fact opens an important way of research in the topic of driving anger, making easier the collection of

data, according to the advantages of Internet-based research explained in the introduction.

Finally, this research has an important limitation. Our sample does not represent the Spanish population drivers accurately, so there is the possibility of sampling bias. This restricts the external validity of the study, because of the limitation of the generalizability of the results to the general population. It would have been desirable to have worked with a stratified sample by age and gender, and that the sample size of each one of the four groups had been more similar and bigger than ours.

5.- References

Alam, I., Khusro, S., Rauf, A., & Zaman, Q. (2014). Conducting surveys and data collection: From traditional to mobile and SMS-based surveys. *Pakistan Journal of Statistics and Operation Research, 10*(2), 169-187.

Bagby, R.M., Ayearst, L.E., Morariu, R.A., Watters, C.; Taylor, G.J. (2014). The Internet administration version of the 20-item Toronto Alexithymia Scale. *Psychological Assessment, 26*(1), 16 – 22.

Bentler, P. M., & Bonnet, D. G. (1980). Significance test and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588-606.

Bollen, K. A. (1989*). Structural equations with latent variables*. New York: Wiley.

Bone, S. A., & Mowen, J. C. (2006). Identifying the traits of aggressive and distracted drivers: A hierarchical trait model approach. *Journal of Consumer Behaviour, 5*(5), 454-464.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing fit. In K. A. Bollen (Ed.), *Testing Structural Equation Models* (pp. 136-162). Sage: Newbury Park, CA.

Buchan, B. D., DeAngelis, D. L., & Levinson, E. M. (2005). A Comparison of the Web-Based and Paper-and-Pencil Versions of the Career Key Interest Inventory With a Sample of University Women. *Journal of Employment Counseling, 42*(1), 39-46.

Buchanan, T. (2000). Potential of the Internet for personality research. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 121-265). San Diego, CA: Academic Press.

Buchanan, T., & Smith, J. L. (1999). Using the Internet for psychological research: Personality testing on the World Wide Web. *British Journal of Psychology, 90*(1), 125-144.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cook, C., Heath, F., Thompson, R. L., & Thompson, B. (2001). Score reliability in Web- or internet-based surveys: Unnumbered graphic rating scales versus Likert-type scales. *Educational and Psychological Measurement, 61*(4), 697-706.

Chang, T.-S. (2005). The Validity and Reliability of Student Ratings: Comparison between Paper-pencil and Online Survey. *Chinese Journal of Psychology, 47*(2), 113-125.

Christie, A., Dagfinrud, H., Dale, O., Schulz, T., & Hagen, K. B. (2014). Collection of patient-reported outcomes; - Text messages on mobile phones provide valid scores and high response rates. *BMC Medical Research Methodology, 14*(1), 1-5.

Dahlen, E. R., Martin, R. C., Ragan, K., & Kuhlman, M. M. (2005). Driving anger, sensation seeking, impulsiveness, and boredom proneness in the prediction of unsafe driving. *Accident Analysis & Prevention, 37*(2), 341-348.

Dahlen, E. R., & Ragan, K. M. (2004). Validation of the Propensity for Angry Driving Scale. *Journal of Safety Research, 35*(5), 557-563.

Davis, R. N. (1999). Web-based administration of a personality questionnaire: Comparison with traditional methods. *Behavior Research Methods, Instruments & Computers, 31*(4), 572-577.

Deffenbacher, J. L., Filetti, L. B., Richards, T. L., Lynch, R. S., & Oetting, E. R. (2003). Characteristics of two groups of angry drivers. *Journal of Counseling Psychology, 50*(2), 123-132.

Deffenbacher, J. L., Lynch, R. S., Deffenbacher, D. M., & Oetting, E. R. (2001). Further evidence of reliability and validity for the Driving Anger Expression Inventory. *Psychological Reports, 89*(3), 535-540.

Deffenbacher, J. L., Lynch, R. S., Filetti, L. B., Dahlen, E. R., & Oetting, E. R. (2003). Anger, aggression, risky behavior, and crash-related outcomes in three groups of drivers. *Behaviour Research and Therapy, 41*(3), 333-349.

Deffenbacher, J. L., Lynch, R. S., Oetting, E. R., & Swaim, R. C. (2002). The Driving Anger Expression Inventory: A measure of how people express their anger on the road. *Behaviour Research and Therapy, 40*(6), 717-737.

Deffenbacher, J. L., Lynch, R. S., Oetting, E. R., & Yingling, D. A. (2001). Driving anger: Correlates and a test of state-trait theory. *Personality and Individual Differences, 31*(8), 1321-1331.

Deffenbacher, J. L., Oetting, E. R., & Lynch, R. S. (1994). Development of a driving anger scale. *Psychological Reports, 74*(1), 83-91.

Deffenbacher, J. L., Richards, T. L., Filetti, L. B., & Lynch, R. S. (2005). Angry Drivers: A Test of State-Trait Theory. *Violence and Victims, 20*(4), 455-469.

Deffenbacher, J. L., White, G. S., & Lynch, R. S. (2004). Evaluation of Two New Scales Assessing Driving Anger: The Driving Anger Expression Inventory and the Driver's Angry Thoughts Questionnaire. *Journal of Psychopathology and Behavioral Assessment, 26*(2), 87-99.

Elander, J., West, R., & French, D. (1993). Behavioral correlates of individual differences in road-traffic crash risk: An examination of methods and findings. *Psychological Bulletin, 113*(2), 279-294.

Epstein, J., Klinkenberg, W. D., Wiley, D., & McKinley, L. (2001). Insuring sample equivalence across Internet and paper-and-pencil assessments. *Computers in Human Behavior, 17*(3), 339-346.

Esiyok, B., Yasak, Y., & Korkusuz, I. (2007). Anger Expression on the Road: Validity and Reliability of the Driving Anger Expression Inventory. *Turkish Journal of Psychiatry, 18*(3), 231-243.

Evans, L. (1991). *Traffic safety and the Driver*. New York: Van Nostrand Reinhold.

Fang, J., Wen, C., & Prybutok, V.R. (2014). An assessment of equivalence between Internet and paper-based surveys: evidence from collectivistic cultures. *Quality and Quantity, 48*(1), 493 – 506.

Field, A. (2005). *Discovering statistics using SPSS*. London: SAGE.

Fouladi, R. T., McCarthy, C. J., & Moller, N. P. (2002). Paper-and-pencil or online? Evaluating mode effects on measures of emotional functioning and attachment. *Assessment, 9*(2), 204-215.

Goehring, J. B. (2000). Aggressive driving: background and overview report. *Environment, Energy and Transportation Program, 2000*.

Hays, S., & McCallum, R. S. (2005). A Comparison of the Pencil-and-Paper and Computer-Administered Minnesota Multiphasic Personality Inventory-Adolescent. *Psychology in the Schools, 42*(6), 605-613.

Herrero-Fernández, D. (2011a). Adaptación psicométrica de la versión reducida del Driving Anger Scale en una muestra española. Diferencias por edad y sexo [Psychometric adaptation of the reduced version of Driving Anger Scale in a Spanish sample. Differences by age and gender]. *Anales de Psicología, 27*(2), 544-549.

Herrero-Fernández, D. (2011b). Psychometric adaptation of the Driving Anger Expression Inventory in a Spanish sample. Differences by age and gender. *Transportation Research Part F: Traffic Psychology and Behaviour, 14*, 324-329.

Hunter, J., Corcoran, K., Leeder, S., & Phelps, K. (2013). Is it time to abandon paper? the use of emails and the Internet for health services research - A cost-

effectiveness and qualitative study. *Journal of Evaluation in Clinical Practice, 19*(5), 855-861.

Lajunen, T., Parker, D., & Stradling, S. G. (1998). Dimensions of driver anger, aggressive and highway code violations and their mediation by safety orientation in UK drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*(2), 107-121.

Lane, S. (2014). Validity evidence based on testing consequences. *Psicothema, 26*(1), 117 – 126.

Leslie, C.S. (2006). *A comparison of self-disclosure measures between Internet and paper-based integrity assessments: Extending research to a job application setting.* Information & Learning, US.

McKee, L. M., & Levinson, E. M. (1990). A review of the computerized version of the Self-Directed Search. *The Career Development Quarterly, 38*(4), 325-333.

Meade, A.W, Michels, L.C., & Lautenschlager, G.J. (2007). Are Internet and Paper-and-Pencil Personality Tests Truly Comparable? An experimental design measurement invariance study. *Organizational Research Methods, 10*(2), 322 – 345.

Moreland, K. L., Zeidner, M., & Most, R. (1992). Computer-assisted psychological assessment. *Psychological testing: An inside view*. (pp. 343-376). Palo Alto, CA US: Consulting Psychologists Press.

Musch, J., Reips, U.-D., & Birnbaum, M. H. (2000). A brief history of Web experimenting. *Psychological experiments on the Internet*. (pp. 61-87). San Diego, CA US: Academic Press.

Oren, C, Kennet-Cohen, T., Turvall, E., & Allalouf, A. (2014). Demosntrating the validity of three general scores of PET in predicting higher education achievement in Israel. *Psicothema, 26*(1), 117 – 126.

Padilla, J.L., & Benítez, I. (2014). Validity evidence base on response processes. *Psicothema, 26*(1), 136 – 144.

Parker, D., Lajunen, T., & Summala, H. (2002). Anger and aggression among drivers in three European countries. *Accident Analysis & Prevention, 34*(2), 229-235.

Parker, D., Reason, J. T., Manstead, A. S. R., & Stradling, S. G. (1995). Driving errors, driving violations and accident involvement. *Ergonomics, 38*(5), 1036-1048.

Pettit, F. A. (1999). Exploring the use of the World Wide Web as a psychology data collection tool. *Computers in Human Behavior, 15*(1), 67-71.

Pinto, D. (2001). *Driving anger, articulated cognitive distortions, cognitive deficiencies, and aggression.* ProQuest Information & Learning, US.

Pomplun, M., & Custer, M. (2005). The Score Comparability Of Computerized And Paper-And-Pencil Formats For K-3 Reading Tests. *Journal of Educational Computing Research, 32*(2), 153-166.

Pomplun, M., Frey, S., & Becker, D. F. (2002). The score equivalence of paper-and-and pencil and computerized versions of a speeded test of reading comprehension. *Educational and Psychological Measurement, 62*(2), 337-354.

Raykov, T., & Marcoulides, G.A. (2006). *A first course in structural equation modelling.* Lawrence Erlbaum Associates: London.

Richman, W. L., Kiesler, S., Weisband, S., & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology, 84*(5), 754-775.

Rios, J., & Wells, C. (2014). Validity evidence based on internal structure. *Psicothema, 26*(1), 108 – 116.

Sampson, J. J. P. (2000). Using the Internet to enhance testing in counseling. *Journal of Counseling & Development, 78*(3), 348-356.

Schwartz, E., & Deffenbacher, J. L. (2002). *Age and Gender Differences in Driving and General Anger and Their Expression*. Paper presented at the 72th Annual Convention of the Rocky Mountain Psychological Association, Park City, Utah.

Sireci, S., & Faulknen-Bond, M. (2014). Validity evidence based on test content. *Psicothema, 26*(1), 100 – 107.

Suarez-Balcazar, Y., Balcazar, F.E., Taylor-Ritzler, T. (2009). Using the Internet to conduct research with culturally diverse populations. Challenges and opportunities. *Cultural Diversity and Ethnic Minority Psychology, 15*(1), 96 – 104.

Sutherland, M. A., Amar, A. F., & Laughon, K. (2013). Who sends the email? using Electronic surveys In violence research. *Western Journal of Emergency Medicine, 14*(4), 363-369.

Underwood, G., Chapman, P., Wright, S., & Crundall, D. (1999). Anger while driving. *Transportation Research Part F: Traffic Psychology and Behaviour*(1), 55-68.

Vispoel, W. P., Boo, J., & Bleiler, T. (2001). Computerized and paper-and-pencil versions of the Rosenberg Self-Esteem Scale: A comparison of psychometric features and respondent preferences. *Educational and Psychological Measurement, 61*(3), 461-474.

Wang, Y.-C., Lee, C.-M., Lew-Ting, C.-Y., Hsiao, C. K., Chen, D.-R., & Chen, W. J. (2005). Survey of substance use among high school students in Taipei: Web-based questionnaire versus paper-and-pencil questionnaire. *Journal of Adolescent Health, 37*(4), 289-295.

Wheaton, B., Muthén, B., Alwil, D., & Summers, G. (1977). Assessing reliability and stability in panel models. In D. R. Heise (Ed.), *Sociological Methodology* (pp. 84-136). San Francisco: Jossey-Bass.

Acknowledgements

Many thanks to Miguel Ángel Vadillo for his comments on a previous draft of this paper. This work was supported by Grant BFI08.292 from the Basque Government.