# TEST-RETEST RELIABILITY OF THE SUSCEPTIBILITY TO DRIVER DISTRACTION QUESTIONNAIRE (SDDQ)

Susana Marulanda
University of Toronto, Department of Mechanical and Industrial Engineering
5 King's College Road, Toronto, ON M5S 3G8, Canada
smarulan@mie.utoronto.ca
(t) 416 978-0881

Huei-Yen Winnie Chen
University of Toronto, Department of Mechanical and Industrial Engineering
5 King's College Road, Toronto, ON M5S 3G8, Canada
hwchen@mie.utoronto.ca
(t) 416 978-0881

Birsen Donmez*
University of Toronto, Department of Mechanical and Industrial Engineering
5 King's College Road, Toronto, ON M5S 3G8, Canada
donmez@mie.utoronto.ca
(t) 416 978-7399
(f) 416 978-7753

* Corresponding author.

Original submission date: August 01, 2014
Revision date: November 15, 2014
Word count: 4,978 + 250 × 3 tables = 5,728

## ABSTRACT

The Susceptibility to Driver Distraction Questionnaire (SDDQ) investigates voluntary and involuntary factors associated with driver distraction. It consists of 39 items in six subscales: (1) self-reported distraction engagement, (2) attitudes towards distractions, (3) perceived control of driving while engaged in distractions, (4, 5) injunctive and descriptive social norms associated with distraction engagement, and (6) susceptibility to involuntary distractions. The test-retest reliability of SDDQ was assessed using a sample of 43 adults, ages 25-39. The mean time between test and retest conditions was approximately 20 days. For subscale averages, test-retest reliability was assessed using intra-class correlation (ICC) statistics; for individual items, it was assessed through weighted kappa statistics. ICC results suggest good to excellent test-retest reliability for subscales of self-reported distraction engagement, attitudes towards distractions, and descriptive social norms. Perceived control of driving while engaged in distractions had fair test-retest reliability, and injunctive norms and susceptibility to involuntary distraction subscales had poor test-retest reliabilities. These last two subscales may have to be redesigned; we provide relevant suggestions in the discussion section. As an additional preliminary analysis, data from a sample of 10 additional participants were used to investigate consistency of responses across longer periods of time. The mean time between test-retest conditions in this sample was approximately 8 months. The findings were in general similar to the main sample. Overall, SDDQ appears to have good test-retest reliability. A larger sample is recommended to further validate these results, in particular across long test-retest periods.

## INTRODUCTION

The diversion of attention away from activities critical for safe driving towards a competing activity is referred to as driver distraction (*1*, *2*). This definition suggests that secondary activities are considered a distraction only when they compete with activities critical for safe driving. A broader definition of driver distraction is used for the purposes of regulation and crash reporting. The U.S. Department of Transportation defines driver distraction as any activity that diverts attention from the primary task of driving (*3*). Some common distractions include cellular phones, which account for 12% of all fatal distraction-based crashes (*4*), in-vehicle assistive and entertainment technologies (*5*), and on-road digital advertising displays (*6*). The U.S. National Highway Traffic Safety Administration (NHTSA) reported that, in 2012, 3,328 people were killed and 421,000 were injured in motor vehicle crashes involving distracted drivers (*4*).

Driver distraction can be the result of a driver's voluntary engagement in a secondary activity, such as talking on the phone. However, certain events or activities can cause a driver to involuntarily divert his attention from driving, such as overhearing a conversation being held among passengers. Research on voluntary engagement in distracting activities while driving shows that such voluntary engagement is usually associated with a positive evaluation of engaging in these activities (*7*), past behavior, confidence in dealing with distraction, perceived risk of distractions, and tendencies towards sensation seeking (*8*). Involuntary distraction, on the other hand, is related to a driver's ability to suppress responses to salient stimuli, which are known to capture attention automatically (*9*). In such a case, a driver may still be distracted by a stimulus or secondary task even though there is no intention to engage in or respond to it.

The Susceptibility to Driver Distraction Questionnaire (SDDQ) (*10, 11*) was developed to measure self-reported engagement in distractions. SDDQ is unique in its aim to distinguish between drivers' susceptibility to voluntary and involuntary distractions. This distinction may significantly improve the understanding of driver distraction and the assessment of motivations behind engagement in secondary tasks. Understanding drivers' susceptibility to each type of distraction can be used to: (1) develop distraction mitigation strategies fitted to individual needs, and (2) assess how new in-vehicle technologies will influence voluntary and involuntary distraction individually and design these technologies in such a way to minimize these influences. Furthermore, SDDQ allows for the examination of whether drivers' judgment of their ability to drive while distracted is based on their ability to effectively evaluate their own attentional abilities. SDDQ is an inexpensive and easy to administer tool that can also facilitate the recruitment of participants from targeted populations, e.g., those who are most susceptible to voluntary distractions, for the evaluation of new in-vehicle technologies or distraction mitigation systems.

To ensure that data collected using SDDQ is meaningful and useful, its validity and reliability must be established. According to Litwin (*12*), validity refers to the degree to which items comprising the questionnaire reflect the constructs that they were designed to measure (e.g., susceptibility to voluntary and involuntary distraction). Several types of validity may be used to assess the performance of a questionnaire: content, criterion, and construct validity. Content validity is established by having subject matter experts review how well the items comprising the questionnaire appear to measure the desired construct. Criterion validity compares the questionnaire's measures with other well established

measures in two forms: (1) concurrent validity, where the questionnaire being examined and the other measures are completed at the same time; and (2) predictive validity, which is concerned with the questionnaire's ability to forecast future responses or behaviors. Finally, construct validity is a measure of how meaningful the questionnaire is for practical use, i.e., generalizable across different settings and times.

Litwin (*12*) defines reliability as the degree to which responses to the survey instrument are reproducible. Reliability is assessed in three forms: test-retest, alternate-form, and internal consistency. Test-retest reliability is a measure of the stability of responses over time, typically by administering the survey at two different periods of time with the same group of respondents. Alternate-form reliability uses different versions (e.g., wording or order of items) of the same questionnaire to assess the same attribute. Internal consistency measures how well a group of items in the questionnaire measures the same construct (*12*).

Since its initial development, SDDQ has been subjected to continual assessments of validity and reliability. To achieve content validity, we began with constructing items based on prevalent distractions: conversations with passengers are reported by NHTSA as the distraction drivers most frequently engage in (*13*); cell phone use and in-vehicle technologies are also distractions identified in NHTSA's crash databases (*4*). Furthermore, we followed the taxonomy from (*14*) to include distractions stemming from different kinds of sources: inside (e.g., in-vehicle technology) and outside (e.g., roadside advertisements and accident scenes) of the vehicle; technology-based (e.g., cell phones) and non-technology based (e.g., passengers); and internal (e.g., daydreaming) and external to the driver. Suggestions brought about by experts in driver distraction, through blind reviews in other publications (*10*, *11*), have also been incorporated in the questionnaire to enhance its content validity.

In an earlier study, we examined the internal consistency of SDDQ and found it to be moderate to high, with Cronbach's alpha ranging from 0.66-0.80 across the different sections of the questionnaire (*10*, *11*). In this study, we also examined the concurrent validity of SDDQ using well-established questionnaire measures of risky driving behaviors, personality traits, and attentional capacities. Findings showed moderate correlations between self-reported engagement and other self-reported measures of unsafe driving behaviors, as assessed by the Manchester Driver Behavior Questionnaire (*15*). Personality traits, such as sensation seeking, measured by the Arnett Inventory of Sensation Seeking (*16*) and impulsiveness, assessed using the Eysenck Impulsivity Questionnaire (*17*), were found to be associated with positive attitudes and beliefs that motivate voluntary engagement in distraction. The study also found that susceptibility to involuntary distraction is related to subjective assessment of cognitive limitations, as measured by the Cognitive Failures Questionnaire (*18*). Overall, these correlations to existing and widely-used questionnaires provided support that SDDQ measures were useful in differentiating between voluntary and involuntary aspects of distraction.

In the current paper, we focus on the test-retest reliability of SDDQ. As mentioned previously, test-retest reliability is concerned with achieving the same results when administering a measure to the same person, in the same way, at different occasions. This form of reliability is useful in ensuring that the instrument provides consistent measures, thus reducing the chance that results may be due to confounding factors (*19*).

## DESCRIPTION OF SDDQ

SDDQ is comprised of 39 items measuring six different constructs or subscales: (1) self-reported distraction engagement, (2) attitudes towards distractions, (3) perceived control of driving while engaged in distractions, (4, 5) injunctive and descriptive social norms associated with distraction engagement, and (6) susceptibility to involuntary distractions (Table 1). Overall, the questionnaire is divided into three major sections as follows:

### Section 1: Engagement in distraction while driving

The first section assesses self-reported frequency of distraction engagement (construct 1) by collecting responses on seven driver distractions: have phone conversations, manually interact with a phone (e.g., sending text messages), adjust the settings of in-vehicle technology (e.g., radio channel or GPS), read roadside advertisements, visually dwell on roadside accident scenes if there are any, chat with passengers if there are any, and daydream. Responses on this section are collected on a 5-point Likert scale comprised of 'never', 'rarely', 'sometimes', 'often', and 'very often'. For scoring purposes, these anchors are assigned points from 1 (never) to 5 (very often) and the points are then averaged across the seven distractions to create an overall section score.

### Section 2: Attitudes and beliefs about voluntary distraction

The second section of the questionnaire investigates facilitators of voluntary distraction through the Theory of Planned Behavior (TPB) (*20*). This section covers constructs 2-5: attitudes, perceived control, and perceived descriptive norms and injunctive norms. Descriptive norms refer to an individual's belief about other peoples' behaviors, while injunctive norms describe the perceived expectations of how an individual ought to behave (*21*). Each construct is probed for the same list of distractions used in Section 1, except for 'daydream', as it is not a distraction that can be voluntarily engaged in by drivers. Responses in this section are collected using a 5-point Likert scale anchored at 'strongly disagree' (1), 'disagree', 'neutral', 'agree', and 'strongly agree' (5). A score for each of the four constructs is calculated by averaging the responses to the six distractions.

### Section 3: Susceptibility to involuntary distraction

The final section of the questionnaire investigates susceptibility to involuntary distraction (construct 6) based on drivers' self-reported ability to suppress stimuli brought about by technologies (i.e., phone and radio), passengers, distractions external to the vehicle, and daydreaming. Distraction in these items is hypothesized to originate from the content of the stimuli (e.g., music or audio alert) rather than the action itself. For example, with respect to the item "While driving, I find it distracting when I listen to music", the action of turning on music is voluntary (i.e., having the radio "on" or "off"), but once the music is being played, the driver may be paying attention to the music involuntarily. Responses for this section measure agreement to relevant statements using a 6-point scale of 'strongly disagree' (1), 'disagree', 'neutral', 'agree', 'strongly agree' (5), and 'never happens'. For scoring purposes, responses across all eight items are averaged excluding responses of 'never happens.'

## TABLE 1 Susceptibility to Driver Distraction Questionnaire (SDDQ)

**(**Headings in brackets (e.g., [Attitudes], [Perceived control]) were not presented to the participants)

| **[Section 1: Distraction Engagement]** | Never | Rarely | Sometimes | Often | Very Often |
|---|---|---|---|---|---|
| ***When driving, I:*** | | | | | |
| a. have phone conversations. | | | | | |
| b. manually interact with a phone (e.g., sending text messages). | | | | | |
| c. adjust the settings of in-vehicle technology (e.g., radio channel or GPS). | | | | | |
| d. read roadside advertisements. | | | | | |
| e. visually dwell on roadside accident scenes if there are any. | | | | | |
| f. chat with passengers if there are any. | | | | | |
| g. daydream. | | | | | |

| **[Section 2: Attitudes and Beliefs about Voluntary Distraction]** | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
|---|---|---|---|---|---|
| **[Attitude]** ***I think it is all right to drive and:*** | | | | | |
| a. have phone conversations. | | | | | |
| b. manually interact with a phone (e.g., sending text messages). | | | | | |
| c. adjust the settings of in-vehicle technology (e.g., radio channel or GPS). | | | | | |
| d. read roadside advertisements. | | | | | |
| e. visually dwell on roadside accident scenes if there are any. | | | | | |
| f. chat with passengers if there are any. | | | | | |
| **[Perceived control]** ***I believe I can drive well even when I:*** | | | | | |
| a. have phone conversations. | | | | | |
| b. manually interact with a phone (e.g., sending text messages). | | | | | |
| c. adjust the settings of in-vehicle technology (e.g., radio channel or GPS). | | | | | |
| d. read roadside advertisements. | | | | | |
| e. visually dwell on roadside accident scenes if there are any. | | | | | |
| f. chat with passengers if there are any. | | | | | |
| **[Perceived social norms 1]** ***Most drivers around me drive and:*** | | | | | |
| a. have phone conversations. | | | | | |
| b. manually interact with a phone (e.g., sending text messages). | | | | | |
| c. adjust the settings of in-vehicle technology (e.g., radio channel or GPS). | | | | | |
| d. read roadside advertisements. | | | | | |
| e. visually dwell on roadside accident scenes if there are any. | | | | | |
| f. chat with passengers if there are any. | | | | | |
| **[Perceived social norms 2]** ***Most people who are important for me think, it is all right for me to drive and:*** | | | | | |
| a. have phone conversations. | | | | | |
| b. manually interact with a phone (e.g., sending text messages). | | | | | |
| c. adjust the settings of in-vehicle technology (e.g., radio channel or GPS). | | | | | |
| d. read roadside advertisements. | | | | | |
| e. visually dwell on roadside accident scenes if there are any. | | | | | |
| f. chat with passengers if there are any. | | | | | |

| **[Section 3: Susceptibility to Involuntary Distraction]** | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree | Never Happens |
|---|---|---|---|---|---|---|
| ***While driving, I find it distracting when*** | | | | | | |
| a. my phone is ringing. | | | | | | |
| b. I receive an audio alert from my phone (e.g., incoming text message). | | | | | | |
| c. I listen to music. | | | | | | |
| d. I listen to talk radio. | | | | | | |
| e. there are roadside advertisements. | | | | | | |
| f. there are roadside accident scenes. | | | | | | |
| g. a passenger speaks to me. | | | | | | |
| h. I daydream. | | | | | | |

## METHODS

### Participants

Data for this paper were collected as part of an ongoing study aimed at further validating SDDQ using performance in a driving simulator and computer-based measures of selective attention and executive function. To be eligible for this study, participants had to be between the ages of 25 and 39 years old, have a valid full driver's license, and have normal or corrected vision.

The sample analyzed in the current paper consists of 25 males and 18 females, a total of 43 individuals. Participants were recruited using online advertisements and posts at local communities. Participants' age ranged from 25 to 39, with a mean age of 29.2 years and a standard deviation of 4.2. They have had their license for an average of 9.03 years (SD=6.03). 35% of participants reported driving under 5,000km, 49% reported driving between 5,001km and 25,000km, and 2% reported driving over 45,001km over the last year; 14% reported not knowing. In addition, on a scale of 1 to 10, with 1 being a very unsafe driver and 10 being a very safe one, 98% of participants rated themselves as a 7 or above (Mean=8.4, SD=0.9).

Ten additional participants, who had been recruited for an earlier study (*10, 11*), for which they completed SDDQ already, were also asked to participate by completing SDDQ for a second time. Although this additional sample was small in size, it gave us the opportunity to conduct preliminary analysis on the reliability of responses in SDDQ across longer periods of time. Data from these participants were analyzed separately.

This additional sample consists of 4 males and 6 females. Age ranged from 25 to 34 (Mean=28.8, SD=3.5). The participants have had their license for an average of 7.4 years (SD=4.7). From these participants, 40% reported driving under 1,600km, 30% reported driving between 1,601 and 8,000km, 20% reported driving between 8,001 and 16,000km, and 1% reported driving over 32,001km over the last year. On the scale of 1-10 measuring how safe of a driver they are, 90% of participants rated themselves as a 7 or above (Mean=7.6, SD=1.6).

### Procedures

Participants completed SDDQ for the first time (test condition) when they filled out an eligibility questionnaire prior to the study, and once again after the completion of the study (retest condition). Average time between test and retest was 19.5 days (range: 0.1-83.2 days, median=9.0 days).

For the 10 participants invited based on their participation in the previous SDDQ study, their original responses to SDDQ were used as the test condition. Data for the retest condition were collected after the completion of the current validation study. For this group, average time between test-retest conditions was 7.9 months (range: 4.2-10.8 months, median=7.8 months).

For both groups, the questionnaire was administered online for both the test and the retest conditions. However, during the test condition, the questionnaire was sent to participants and was completed outside of the lab, whereas for the retest condition, the questionnaire was administered in a lab setting.

## RESULTS

Test-retest reliability analyses were conducted in SAS version 9.3. Intra-class correlation coefficients (ICC) (Type 1,1) were computed to assess the test-retest reliability at the subscale level (*22*) (Table 2). ICC reliability values can range from 0 to 1: values less than 0.4 represent poor agreement, those between 0.4 and 0.6 are fair, values between 0.6 and 0.75 are good, and those greater than 0.75 are considered excellent (*23*). In addition, weighted kappa statistics (*24*) were used to assess the reliability of test-retest pairs for individual items (Table 3). Agreement ratings suggested by Landis & Koch (*25*) were used to interpret kappa values: less than 0 are poor, between 0 and 0.2 are slight, between 0.21 and 0.4 are fair, between 0.41 and 0.6 are moderate, between 0.61 and 0.8 are substantial, and between 0.8 and 1 almost perfect.

### Subscale Reliability

ICC (Type 1,1) statistics were calculated for all subscales (Table 2). For the sample of 43 participants, who were retested within approximately 20 days on the average, good to excellent test-retest reliability was demonstrated for most subscales of SDDQ. Only two of the subscales, i.e., injunctive social norms and involuntary distraction, were found to have poor ICCs (i.e., 0.35 and 0.37). For the additional 10 participants, who were retested after several months, ICCs for descriptive and injunctive social norm subscales were poor. For subscales of self-reported distraction engagement and attitudes toward distractions, ICCs were good to excellent. Finally, ICCs of the perceived control subscale were fair in both samples.

**TABLE 2 Intra-class correlation coefficient (Type 1,1) for all subscales of SDDQ**

| | **Test-retest period** | |
|---|---|---|
| **SDDQ Section** | **~20 days (n=43)** | **~8 months (n=10)** |
| Engagement | 0.77*** | 0.77*** |
| Attitudes | 0.74** | 0.73** |
| Perceived control | 0.59* | 0.59* |
| Social norms | | |
| *Descriptive* | 0.63** | 0.36 |
| *Injunctive* | 0.37 | 0.00 |
| Involuntary distraction | 0.35 | 0.61** |

*fair, **good, ***excellent based on recommended values (*23*)

### Item Reliability

Due to the poor ICCs found in the sample of 43 participants for the injunctive social norms and involuntary distraction subscales, items surveyed for these two subscales were investigated separately using weighted kappa statistics. Weighted kappa statistics were not calculated for the sample of additional 10 participants, due to the small sample size. In general, the reliability of the individual items comprising the injunctive norms and involuntary distraction subscales was between fair and substantial (weighted kappa ranged between 0.23 and 0.64; Table 3).

**TABLE 3 Kappa statistics and 95% confidence intervals (CI) for injunctive norms and involuntary distraction subscales**

| Item | Weighted Kappa | 95% CI | Agreement Strength |
|---|---|---|---|
| ***Injunctive Social Norms*** | | | |
| Phone conversations | 0.37 | 0.09, 0.66 | Fair |
| Manual interaction with phone | 0.44 | 0.13, 0.75 | Moderate |
| Adjust in-vehicle technology | 0.36 | 0.11, 0.62 | Fair |
| Roadside advertisements | 0.23 | -0.04, 0.51 | Fair |
| Roadside accident scenes | 0.29 | -0.01, 0.59 | Fair |
| Chat with passengers | 0.43 | 0.16, 0.70 | Moderate |
| ***Involuntary Distraction*** | | | |
| Phone ringing | 0.44 | 0.16, 0.72 | Moderate |
| Phone alert | 0.57 | 0.33, 0.82 | Moderate |
| Listen to music | 0.49 | 0.26, 0.71 | Moderate |
| Listen to talk radio | 0.64 | 0.36, 0.91 | Substantial |
| Roadside advertisements | 0.47 | 0.19, 0.75 | Moderate |
| Roadside accident scenes | 0.52 | 0.28, 0.75 | Moderate |
| Passengers speaking | 0.40 | 0.12, 0.68 | Fair |
| Daydream | 0.75 | 0.58, 0.91 | Substantial |

## DISCUSSION

This study examined the test-retest reliability of SDDQ, a recently developed questionnaire designed to assess voluntary and involuntary facilitators associated with driver distraction. Forty-three participants were retested within approximately 20 days from their initial testing. In this sample, drivers responded consistently to the majority of SDDQ subscales. The test-retest reliability for the following subscales, as measured by ICCs, ranged between good and excellent: self-reported distraction engagement, attitudes towards distractions, and descriptive social norms. An additional sample of 10 drivers were also tested and retested, but their test-retest period was in the order of months; thus this additional sample provided a preliminary analysis on longer-term test-retest reliability. ICCs for self-reported distraction engagement and attitudes towards distractions for this additional sample also ranged from good to excellent, despite the fact that there was an average of 8 months between the test and the retest conditions. Overall, the results of the additional participants were similar to the results obtained from the main sample, with the exception of ICCs for descriptive norms and involuntary distraction subscales. The small sample size for the additional sample, however, is a big limitation and these findings have to be confirmed with a larger sample size. For example, reliability of responses for involuntary distraction increased with the longer test-retest period, which is unexpected. The decrease in reliability for the descriptive norms, however, can in part be explained by the respondents having to average their perceptions of other drivers' behavior over a longer period of time. The respondents may have based their retest answers on recent events (e.g., a penalty increase for texting while driving), which may have differed substantially from those at the initial testing period.

In the main sample of 43 participants (shorter test-retest period), injunctive norms and involuntary distraction subscales had poor ICCs, indicating potential issues for these two subscales. In general, responses in the social norms subscales may be more variable due to the difficulty associated with making inferences about others' behaviors compared to making judgments about oneself. However, this effect may be less pronounced for descriptive norms compared to injunctive norms, as behavior is more transparent than opinions, as the former is observable. To examine the individual items within the injunctive norms subscale as well as the involuntary distraction subscale, weighted kappa statistics were used. The individual items in these subscales had fair to substantial test-retest reliability, therefore they did not appear to be problematic overall. However, the poor ICC value in the injunctive norms subscale appeared to be driven by items corresponding to 'roadside advertisements' and 'roadside accident scenes', which had the lowest weighted kappa values.

These findings from weighted kappa values suggest that some distractions used in the injunctive norms subscale may not be appropriate for measuring this construct. It is possible that some of the distractions may not have strong social norms attached to them, as may be the case for items such as 'read roadside advertisements' or 'visually dwell on roadside accident scenes'. Without a firm belief of society's approval or disapproval associated with engaging in these particular distractions, drivers' opinions about how they ought to behave are more likely to change over time, resulting in response inconsistencies. Also, due to a lack of societal norms regarding these items as well as their constant presence in the driving environment, they may not be considered distractions in the same manner as other non-driving-related tasks are (e.g., texting while driving). As a result, drivers may be less aware of their own engagement, and consequently their memory of their engagement in these distractions may not be accurate.

Furthermore, some of the distractions may not have been specific enough to elicit the established social norms. For example, campaigns against driving and texting have become popular, yet this awareness was only reflected in the questionnaire as part of 'manually interacting with a phone (e.g., sending text messages).' As the wording of the item may elicit responses for activities other than the example provided, e.g., using a cell phone as a musical device, respondents may have been required to narrow down different activities to a single response. As a result, respondents may have based their answer on those activities in which they have engaged in most recently rather than their experiences with cell phones as a whole.

The poor ICC value for the involuntary distraction subscale also suggests issues with this subscale. It is hypothesized that a lack of context for the distractions measured in this subscale is responsible for the poor test-retest reliability. Context is an important factor in understanding susceptibility to involuntary driver distraction, as the perception of the distractibility of a stimulus may change depending on the environment. For example, a cell phone ringing may be perceived as being more distracting when the driver is in a complex urban environment, which requires considerable attention, compared to when she is in a rural environment that demands less attention. Similar to the injunctive norms subscale, the involuntary distraction subscale may force respondents to generalize behaviors, in this case, across different environments. Consequently, reliability of responses may decrease, as drivers may base their responses on recent events rather than on their overall experiences.

Overall, findings from this study suggest that a revision of the injunctive norms and involuntary distraction subscales is warranted. Based on the results of the current paper, it would be most beneficial to remove the items "visually dwell on roadside advertisements if

there are any" and "read roadside advertisements" from the descriptive and injunctive social norms subscales, and to include only the distractions for which social norms are well established. In addition, it would be best to avoid lumping different distractions in a single item, such as "manually interact with a phone (e.g., sending text messages)," and instead have individual items for different distractions, e.g., "key in text messages" and "read emails on your phone." Further, including context in the questions, such as time of day and driving environment, could help respondents narrow down their experiences with distraction engagement and thus report their behaviors more accurately. The context can be specified by introducing variables of location and time of day in individual questions, or by defining a scenario at the beginning of the questionnaire and instructing participants to respond to all questions based on their experiences in that particular scenario. However, it is important to note that a tradeoff exists between providing specific questions and being able to capture a wide range of distractions without significantly lengthening the questionnaire.

The results of this study are limited by a small sample size. Although over 500 people have completed SDDQ since its development, only 53 drivers have retaken the test, thus providing a limited sample for assessing test-retest reliability. In addition, recruitment of the sample through online posts may have introduced self-selection bias, as it is possible that people who chose to take part in the study may have differed in their motivations or characteristics from those who chose not to participate. It is also possible that the participants who completed the driving simulator study may have been biased in their retest responses, as they were presented with secondary tasks in the simulator. Therefore, the activity conducted by the sample of 43 participants between the test and retest of the questionnaire is another limitation of the current study. A final limitation is the range of test-retest periods employed on this sample of 43 participants (0.1 to 83.2 days). Due to sample size considerations, we avoided further breaking down our sample to investigate if this relatively large range had an effect on reliability. However, it should be noted that participants might have been able to remember their previous responses, in particular if they were retested within the same day. In fact, there were three participants who were retested within the same day. We conducted a sensitivity analysis by removing these three participants from our sample and re-analyzing the remaining 40 participants' data. No differences were found. Future studies should collect data in shorter ranges with more participants within each range.

It should be noted that due to its self-report nature, validity of SDDQ is limited to respondents' introspective ability, their understanding of the rating scales, and their social and memory biases. Most predominantly, as a self-reported measure of aberrant behavior, socially desirable responding is likely to bias respondents, potentially leading to under-reporting in some SDDQ items. Hence, some items, especially those regarding the use of cell phones, can be expected to be more sensitive to social desirability bias.

Despite some limitations, the analyses presented in this paper demonstrate that SDDQ has good test-retest reliability for most of its subscales. In addition, these results also provide valuable insight for a revision of SDDQ. We stress that this study is an early step in the validation of SDDQ. Studies investigating the relationship between SDDQ's measures of involuntary distraction and objective measures of distractibility (i.e., cognitive measures of attention) are underway to ensure the predictive validity of SDDQ (*26*). Furthermore, the predictive validity of SDDQ will be established using performance in a driving simulator with a self-paced secondary task, as well as in the presence of potentially involuntary

distractions. Data collected from these analyses will be subjected to rigorous analysis to establish the overall validity and reliability of the scale.

## ACKNOWLEDGMENTS

The funding for this work was provided by the Toyota Collaborative Safety Research Center (CSRC) and Auto 21 Network of Centres of Excellence. Many thanks to James Foley, Kazu Ebe, and Chuck Gulash from Toyota CSRC for their feedback on this paper as well as providing directions for our research in general.

## REFERENCES


1. Foley, J., R. Young, L. Angell, and J. E. Domeyer. Towards Operationalizing Driver Distraction. In *Proceedings of the 7th International Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, Bolton Landing, NY, 2013.
2. Lee, J. D., K. L. Young, and M. A. Regan. Defining driver distraction. In M.A., Regan, J.D. Lee and K. Young (Eds.) *Driver Distraction: Theory, Effects, and Mitigation* (pp.31–40). CRC Press: Boca Raton, FL, 2008.
3. National Highway Traffic Safety Administration. *What is Distracted Driving? Official US Government Website for Distracted Driving.* www.distraction.gov/content/get-the-facts/facts-and-statistics.html Accessed May 1, 2014.
4. National Highway Traffic Safety Administration. *Distracted Driving 2012.* No. DOT HS 812 012. US Department of Transportation, 2014.
5. Transport Canada. *Strategies for Reducing Driver Distraction from In-Vehicle Telematics Devices: A Discussion Document.* TP 14133E. Transport Canada, 2003.
6. Rempel, G., M. Moshiri, and J. Montufar. Considerations for Assessing the Road Safety Impact of Digital and Projected Advertising Displays in Canada. In *Proccedings of the 2013 Annual Conference of the Transporation Association of Canada*, Winnipeg, Manitoba, 2013.
7. Walsh, S. P., K. M. White, M. K. Hyde, and B. Watson. Dialling and Driving: Factors Influencing Intentions to Use a Mobile Phone While Driving. *Accident Analysis and Prevention*, Vol. 40, No. 6, 2008, pp. 1893–900.
8. Horrey, W. J., and M. F. Lesch. Factors Related to Drivers' Self-Reported Willingness to Engage in Distracting In-Vehicle Activities. In *Proceedings of the Human Factors Ergonomics Society 52nd Annual Meeting*, New York, NY, pp. 1546–1550, 2008.
9. Franconeri, S., and D. Simons. Moving and Looming Stimuli Capture Attention. *Perception and Psychophysics*, Vol. 65, No. 7, 2003, pp. 999–1010.
10. Feng, J., S. Marulanda, and B. Donmez. Susceptibility to Driver Distraction Questionnaire (SDDQ): Development and Relation to Self-Reported Measures. In *Proceedings of Transportation Research Board – 93rd Annual Meeting* (14-3009), Washington, D.C., 2014.
11. Feng, J., S. Marulanda, and B. Donmez. Susceptibility to Driver Distraction Questionnaire (SDDQ): Development and Relation to Self-Reported Measures. *Transportation Research Record*, in press.
12. Litwin, M.S. *How to Measure Survey Reliability and Validity*. SAGE Publications, Inc., 1995.
13. Schroeder, P., M. Meyers, and L. Kostyniuk. *National Survey on Distracted Driving, Attitudes, and Behaviors.* No. DOT HS 811 729. National Highway Traffic Safety Administration, 2013.
14. Regan, M. A., K. L. Young, J. D. Lee, and C. P. Gordon. Sources of driver distraction. In M.A., Regan, J.D. Lee, and K. Young (Eds.) *Driver Distraction: Theory, Effects, and Mitigation* (pp.249-279). CRC Press: Boca Raton, FL, 2008.
15. Lajunen, T., D. Parker, and H. Summala. The Manchester Driver Behaviour Questionnaire: A Cross-Cultural Study. *Accident Analysis and Prevention*, Vol. 36, No. 2, 2004, pp. 231–238.

16. Arnett, J. J. Sensation Seeking, Aggressiveness, and Adolescent Reckless Behavior. *Personality and Individual Differences*, Vol. 20, No. 6, 1996, pp. 693–702.
17. Eysenck, S., P. Pearson, G. Easting, and F. Allsopp. Age Norms for Impulsiveness, Venturesomeness and Empathy in Adults. *Personality and Individual Differences*, Vol. 6, No. 5, 1985, pp. 613–619.
18. Broadbent, D. E., P. F. Cooper, P. FitzGerald, and K. R. Parkes. The Cognitive Failures Questionnaire (CFQ) and its Correlates. *British Journal of Clinical Psychology*, Vol. 21 (Pt 1), 1982, pp. 1–16.
19. Shrout, P., R. Spitzer, and J. Fleiss. Quantification of Agreement in Psychiatric Diagnosis Revisited. *Archives of General Psychiatry*, Vol. 44, 1987, pp. 172–177.
20. Ajzen, I. The Theory of Planned Behavior. *Organizational Behavior and Human Decision Process*, Vol. 50, No. 2, 1991, pp. 179–211.
21. Cestac, J., F. Paran, and P. Delhomme. Drive as I Say, Not as I Drive: Influence of Injunctive and Descriptive Norms on Speeding Intentions Among Young Drivers. In *Transportation Research Record Part F: Traffic Psychology and Behavior*, Vol. 23, 2014, pp. 44–56.
22. Weir, J. Quantifying Test-Retest Reliability Using the Intraclass Correlation Coefficient and the SEM. *The Journal of Strength and Conditioning Research*, Vol. 19, No. 1, 2005, pp. 231–240.
23. Fleiss, J. L., B. Levin, and M. C. Paik. *Statistical Methods for Rates and Proportions*. John Wiley & Sons, Inc., New York, 1981.
24. Cohen, J. Weighted kappa: Nominal Scale Agreement with Provision for Scaled Disagreement or Partial Credit. *Psychological Bulletin*, Vol. 70, No. 4, 1968, pp. 213–20.
25. Landis, J. R., and G. G. Koch. The Measurement of Observer Agreement for Categorical Data. *Biometrics*, Vol. 33, No. 1, 1977, pp. 159–174.
26. Hoekstra-Atwood, L., and H.Y.W. Chen. Measuring Inhibitory Control in Driver Distraction. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, ACM, 2014.