

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
COMISSÃO DE GRADUAÇÃO DO CURSO DE ENGENHARIA CIVIL**

Lucas Vanzella Rocha

**ANÁLISE DE CLUSTERS A PARTIR DE DADOS
EXPERIMENTAIS E AUTORRELATADOS PARA
IDENTIFICAÇÃO DE PERFIS DE CONDUTORES**

Porto Alegre
julho, 2019

LUCAS VANZELLA ROCHA

**ANÁLISE DE CLUSTERS A PARTIR DE DADOS
EXPERIMENTAIS E AUTORRELATADOS PARA
IDENTIFICAÇÃO DE PERFIS DE CONDUTORES**

Trabalho de Conclusão de Curso apresentado à Comissão de
Graduação do Curso de Engenharia Civil da Escola de Engenharia
da Universidade Federal do Rio Grande do Sul, como parte dos
requisitos para obtenção do título de Engenheiro Civil

Orientadora: Prof^a. Christine Tessele Nodari

Porto Alegre

Julho, 2019

LUCAS VANZELLA ROCHA

**ANÁLISE DE CLUSTERS A PARTIR DE DADOS
EXPERIMENTAIS E AUTORRELATADOS PARA
IDENTIFICAÇÃO DE PERFIS DE CONDUTORES**

Este Trabalho de Diplomação foi julgado adequado como pré-requisito para a obtenção do título de ENGENHEIRO CIVIL e aprovado em sua forma final pela Banca Examinadora, pela Professora Orientadora e pela Comissão de Graduação do Curso de Engenharia Civil da Universidade Federal do Rio Grande do Sul.

Porto Alegre, julho de 2019

BANCA EXAMINADORA

Prof^ª. Christine Tessele Nodari (UFRGS)

Dra. pela Universidade Federal do Rio Grande do Sul
Orientadora

Tânia Batistela Torres (UFRGS)

Me. pela Universidade Federal do Rio Grande do Sul

Prof^ª. Raquel da Fonseca Holz (UFPEL)

Dra. Pela Universidade Federal do Rio Grande do Sul

Dedico este trabalho aos meus pais, Andre e Inês, pela paciência e dedicação que demonstraram durante o período de graduação.

AGRADECIMENTOS

Agradeço a professora Christine Tessele Nodari pelo conhecimento transmitido e por toda a atenção dedicada para a elaboração deste trabalho.

Agradeço a Engenheira Tânia Batistela Torres e a professora Raquel Holz pelas considerações relevantes para uma futura publicação.

Agradeço aos meus pais, Andre e Inês, por sempre estarem junto comigo, me dando todo o suporte necessário para concluir o curso.

Agradeço a minha namorada, Paulinha, pelo amor e apoio incondicional, e por me ajudar em tudo que precisei durante a execução deste trabalho.

Agradeço aos diversos amigos e pessoas especiais que conheci durante o período do curso, que ajudaram a tornar a trajetória inesquecível.

Ideias e somente ideias podem iluminar a escuridão.

Ludwig Von Mises

RESUMO

Este trabalho teve o objetivo de identificar perfis de condutores através de um experimento em simulador de condução imersivo, contando com a participação de 73 voluntários de gêneros e idades variados, com tempo mínimo de carteira de habilitação de 1 ano. Os voluntários responderam um total de 32 questões, sobre suas características como condutor e a experiência com o simulador. Foi utilizado um Agrupamento Hierárquico para a formação de 3 *clusters*, as características autorrelatadas juntamente com os dados de velocidade média obtidos no experimento foram comparados entre os *clusters*. Resultados mostram um grupo mais inexperiente, com média de idade de 26 anos, e uma frequência ao volante significativamente menor que os demais. O grupo com média de idade de 32 anos relatou a experiência mais positiva em relação ao simulador, enquanto o *cluster* formado por indivíduos mais maduros e de maioria do sexo masculino relatou a experiência mais negativa em relação à imersão e envolvimento com a simulação.

Palavras-chave: simulador de direção, segurança viária.
Análise Fatorial Exploratória. Análise de *Clusters*.

LISTA DE FIGURAS

Figura 1 – Simulador de Condução Imersivo	19
Figura 2 – Gráfico de Escarpa	26
Figura 3 – Dendograma de formação dos <i>clusters</i>	26
Figura 4 – Médias dos <i>clusters</i> para os 5 Fatores	30
Figura 5 – Pontuação média para o Fator 1 – “Imersão”	31
Figura 6 – Pontuação média para o Fator 2 – “Prática”	32
Figura 7 – Pontuação média para o Fator 3 – “Qualidade do Equipamento”	33
Figura 8 – Pontuação média para o Fator 4 – “Interação com a Realidade Virtual”	34
Figura 9 – Pontuação média para o Fator 5 – “Envolvimento com a Simulação	34

LISTA DE TABELAS

Tabela 1 – Questionário de perfil do condutor	20
Tabela 2 – Questionário de experiência no simulador	21
Tabela 3 – Fatores extraídos	25
Tabela 4 – Distribuição de indivíduos por <i>cluster</i>	27
Tabela 5 – <i>Clusters</i> por sexo	27
Tabela 6 – <i>Clusters</i> por idade	28
Tabela 7 – <i>Clusters</i> por tempo de carteira de habilitação	28
Tabela 8 – <i>Clusters</i> por “Você se considera um condutor experiente ou inexperiente	29
Tabela 9 – <i>Clusters</i> por velocidade média	30
Tabela 10 – Quadro síntese dos <i>clusters</i> e suas características	35

LISTA DE SIGLAS

AFE – Análise Fatorial Exploratória

KMO – Kayser-Meyer-Olkin

ANOVA – Análise Univariada de Variância

DMS – Mínima Diferença Significativa

LISTA DE SÍMBOLOS

χ^2 – Teste Qui-Quadrado.

SUMÁRIO

1 INTRODUÇÃO	14
2 REVISÃO BIBLIOGRÁFICA	15
2.1 FATORES HUMANOS INFLUENCIADORES DA CONDUÇÃO	15
2.2 USO DE ANÁLISE DE AGRUPAMENTO NA MELHORIA DA SEGURANÇA VIÁRIA	17
2.3 SIMULADORES DE DIREÇÃO COMO FERRAMENTA DE PESQUISA	18
3 MÉTODO	18
3.1 SIMULADOR DE DIREÇÃO	19
3.2 CENÁRIO DE ESTUDO	19
3.3 DADOS DA ANÁLISE	20
3.4 PROCEDIMENTO DA ANÁLISE	22
3.4.1 ANÁLISE FATORIAL EXPLORATÓRIA	22
3.4.2 ANÁLISE DE <i>CLUSTERS</i>	23
4 RESULTADOS	25
4.1 ANÁLISE FATORIAL EXPLORATÓRIA	25
4.2 ANÁLISE DE <i>CLUSTERS</i>	26
4.2.1 PERFIS DOS CLUSTERS	27
4.3 ANÁLISE DOS <i>CLUSTERS</i> PELOS FATORES EXTRAÍDOS	30
4.3.1 FATOR IMERSÃO	31
4.3.2 FATOR PRÁTICA	32
4.3.3 FATOR QUALIDADE DO EQUIPAMENTO	33
4.3.4 FATOR INTERAÇÃO COM A REALIDADE VIRTUAL	33
4.3.5 FATOR ENVOLVIMENTO COM A SIMULAÇÃO	34
5 CONSIDERAÇÕES FINAIS	35
REFERÊNCIAS	37

ANÁLISE DE CLUSTERS A PARTIR DE DADOS EXPERIMENTAIS E AUTORRELATADOS PARA IDENTIFICAÇÃO DE PERFIS DE CONDUTORES

Lucas Vanzella Rocha
Christine Tessele Nodari

Universidade Federal do Rio Grande do Sul
Escola de Engenharia

RESUMO

Este trabalho teve o objetivo de identificar perfis de condutores através de um experimento em simulador de condução imersivo, contando com a participação de 73 voluntários de gêneros e idades variados, com tempo mínimo de carteira de habilitação de 1 ano. Os voluntários responderam um total de 32 questões, sobre suas características como condutor e a experiência com o simulador. Foi utilizado um Agrupamento Hierárquico para a formação de 3 *clusters*, as características autorrelatadas juntamente com os dados de velocidade média obtidos no experimento foram comparados entre os *clusters*. Resultados mostram um grupo mais inexperiente, com média de idade de 26 anos, e uma frequência ao volante significativamente menor que os demais. O grupo com média de idade de 32 anos relatou a experiência mais positiva em relação ao simulador, enquanto o *cluster* formado por indivíduos mais maduros e de maioria do sexo masculino relatou a experiência mais negativa em relação à imersão e envolvimento com a simulação.

ABSTRACT

This paper had the objective to identify driver profiles through a simulator experiment, with the participation of 73 volunteers of different genders and ages, and a minimum period of 1 year with a driver's license. The volunteers answered a total of 32 questions, about their characteristics as drivers and the experience with the simulator. Hierarchical clustering was used and resulted in the formation of 3 clusters, the data used to compare the clusters consisted of their self-reported characteristics and the average individual speed obtained in the experiment. Results showed a more inexperienced group than the others, mean age of 26 years, and habit of driving significantly lower. The cluster with a mean age of 32 years reported the most positive experience in relation to the simulation, while the cluster composed of more mature individuals on average and a male majority, reported a more negative experience in general, particularly in relation to immersion and involvement with the simulation.

1. INTRODUÇÃO

Sendo a oitava causa de mortes no mundo, os acidentes de trânsito matam cerca de 1,35 milhão de pessoas por ano, além de serem a maior causa de morte de pessoas com idade entre 5 e 29 anos (WHO, 2018). No Brasil, são cerca de 43 mil mortes por ano (Carvalho, 2016), o que o coloca como o quinto país com mais vítimas fatais decorrentes de acidentes de trânsito (WHO, 2018). Essas estatísticas alarmantes levaram a Organização das Nações Unidas a declarar a década de 2011 a 2020 como a Década de Ação Para Segurança Viária, convocando todos os seus apoiadores para um esforço mundial, com medidas que visam minimizar essas estatísticas.

Diante desses números, torna-se imprescindível que ocorram estudos buscando a identificação e o entendimento das causas dos acidentes em relação a diversos fatores, como as características da via, dos veículos, do clima e, principalmente da natureza humana, já que

diversos estudos já revelaram o fator humano como um importante fator contribuinte de acidentes com veículos leves (Perchonok, 1972; Treat et al., 1979; Reason et al., 1990). Portanto, é de fundamental importância que se entenda o comportamento de condução dos motoristas e quais traços de sua natureza podem se tornar causadores de acidentes.

O estudo do comportamento de condução em ambiente real traz muitos desafios, relacionados tanto à disponibilidade de operação da via, quanto aos custos decorrentes de sua utilização e até mesmo da segurança do condutor a ser analisado. É por isso que ferramentas computacionais devem ser utilizadas para o estudo desse comportamento, e o simulador de direção oferece a possibilidade de realizar experimentos em ambientes controlados sem os riscos que derivariam da utilização da via, permitindo uma análise dinâmica de diversos condutores frente à uma mesma situação e uma análise do mesmo condutor perante diversas situações.

Este artigo tem como objetivo realizar uma identificação de perfis de condutores através da formação de *clusters*, em que indivíduos são agrupados a partir de medições em simulador de condução imersivo e dados autorrelatados em dois questionários diferentes, o primeiro envolvendo questões de perfil e o segundo sobre a experiência do condutor com o simulador.

2. REVISÃO BIBLIOGRÁFICA

Esta seção tem como objetivo destacar o embasamento teórico a respeito do tema deste artigo e está dividida em três subseções. A primeira aborda a importância de conhecer os fatores humanos influenciadores da condução segura, principalmente aqueles ligados à personalidade, percepção de risco e atitudes que se traduzem no comportamento do condutor. A segunda busca evidenciar estudos passados que utilizaram uma metodologia similar à deste artigo e a terceira é uma síntese do uso de simuladores de condução imersivos em estudos de segurança viária.

2.1. Fatores Humanos Influenciadores da Condução

As causas de acidentes de veículos motorizados são complexas e estão relacionadas, em boa parte, às características do condutor, como idade, experiência e gênero. Uma tendência observada em todo o mundo é o envolvimento de jovens em acidentes de trânsito, sendo a maior causa de morte de pessoas com idade entre 5 e 29 anos no mundo (WHO, 2018). McCartt et al. (2003) verificaram a inexperiência como uma das maiores causas de infrações e colisões, concluindo que a incidência de acidentes era mais alta para motoristas em seu primeiro mês após retirar a carteira de habilitação, ou em seus primeiros 800 quilômetros dirigidos.

O envolvimento desproporcional do sexo masculino também é observado nestas estatísticas, estudos mostraram a direção perigosa com uma maior frequência entre homens jovens (Harré et al., 2000; Begg e Langley., 2001). As estatísticas brasileiras reforçam esses números, já que homens são as maiores vítimas de acidentes fatais no Brasil, cerca de 80% (Carvalho, 2016). Navas et. al (2019) conduziram um estudo com a participação de 506 estudantes universitários, com idades variando entre 18 e 30 anos e chegaram à conclusão que homens mostraram uma menor percepção de risco que mulheres e uma maior frequência de condução sob a influência de álcool, assim como uma maior “busca por sensações fortes”, aquilo que Zuckerman (1979) definiu como *sensation seeking*. Outras pesquisas chegaram a conclusões distintas, encontrando taxas mais elevadas de envolvimento em acidentes entre idosos, fazendo uma média por quilômetro dirigido, do que em outras faixas etárias (Retchin e Anapolle, 1993; Massie, 1995).

Diante desses resultados, emergem questões sobre as distinções existentes no comportamento ao volante entre homens e mulheres, ao estudar tais diferenças algumas abordagens se tornam frequentes, como as avaliações das diferenças de variações de humor entre ambos. Dahlen e Ragan (2004), em estudo envolvendo 242 estudantes universitários, apontaram uma maior incidência em expressões de agressão física por indivíduos do sexo masculino, enquanto mulheres tiveram um resultado mais expressivo em agressões verbais. O estudo foi desenvolvido com o questionário *Propensity for Driving Angry Scale*, criado com o objetivo de identificar indivíduos mais propensos a se tornarem irritados enquanto dirigem (DePasquale et al., 2001).

Seguindo a linha de raciocínio estabelecida por WHO (2017), em que é feita uma separação entre as definições de sexo e gênero, alguns autores sugerem que não é o sexo que está ligado à uma maior expressão de raiva no trânsito, mas sim os demais atributos tradicionalmente ligados a ele, sendo masculinidade e feminilidade os termos utilizados para categorizar estes atributos.

Sullman et al.(2017) conduziram um estudo com amostragem de 378 condutores franceses, onde 38% eram homens, e através de uma Análise Fatorial Confirmatória chegaram à conclusão que a masculinidade estava associada à uma maior expressão de raiva no trânsito, inclusive utilizando o veículo como ferramenta de expressão de raiva, apesar de não encontrar correlação entre estes fatores e uma maior incidência de acidentes. A relação entre masculinidade e comportamento agressivo no trânsito também foi encontrada por Özkan e Lajunen (2005) que sugerem uma maior incidência em direção de risco e comportamento agressivo junto ao comportamento masculino. A masculinidade, quando aliada à juventude

também foi associada à condução em maiores velocidades (Schmid Mast et al., 2008).

2.2. Uso de Análise de Agrupamentos na Melhoria da Segurança Viária

Análise de agrupamentos, ou *clustering*, é um conjunto de técnicas exploratórias de análise multivariada de dados, onde o principal objetivo é agrupar objetos similares a partir de suas características (Hair, 2005). Por serem técnicas versáteis, elas são utilizadas em pesquisas de diversas áreas, inclusive em estudos diversificados de segurança viária, como relatado a seguir.

Musicant et al. (2019) agruparam 25 homens, com idades entre 24 e 39 anos, em dois *clusters* de acordo com suas respostas fisiológicas, como frequência cardíaca e variabilidade da frequência cardíaca, durante diferentes processos de frenagem de um veículo como forma de associar a intensidade da desaceleração às respostas involuntárias do corpo. Chegando à conclusão que os indivíduos tiveram respostas suficientemente diferentes a ponto de serem agrupados de formas distintas, porém suas intensidades de frenagem não tiveram diferenças significativas.

Na Noruega, Oltedal e Rundmo (2007) examinaram grupos de indivíduos com traços de personalidade e características culturais similares para analisar as diferenças em percepção de risco entre os *clusters* formados, encontrando diferenças significativas entre os diferentes agrupamentos. Uma abordagem similar foi utilizada por Kaçan et al. (2019), como forma de analisar as similaridades e diferenças no comportamento de motoristas com origem em cinco países do leste europeu.

Lucidi et al. (2010) em um estudo desenvolvido com 1008 estudantes italianos com idades entre 18 e 23 anos, foram capazes de identificar três *clusters* com comportamentos bastante distintos. O primeiro *cluster* foi formado em sua maioria por homens, cerca de 75%, e estava associado a uma baixa responsabilidade no trânsito, pouco respeito com as leis e baixa ansiedade, motoristas desse subgrupo receberam mais multas e se envolveram em mais acidentes que os indivíduos dos outros grupos. O segundo foi caracterizado por escores altos em ansiedade e direção com comportamento de raiva, o que indicaria instabilidade emocional, enquanto o terceiro grupo teve escores baixos em ansiedade e hostilidade, indicando motoristas mais calmos. Resultados similares foram encontrados por Ulleberg (2001), onde um subgrupo formado em sua maioria por homens foi identificado com maior envolvimento em acidentes, baixo nível de ansiedade e altos níveis de *sensation-seeking*.

Na região metropolitana de Porto Alegre, Caleffi et al. (2016) aplicaram métodos de agrupamento para caracterizar o comportamento do tráfego em trecho da rodovia BR-290/RS

em relação a diferentes condições meteorológicas, onde reconheceram diferentes padrões de comportamento sob precipitações. Os pesquisadores também identificaram que chuvas leves levaram à redução das médias de velocidade em 10 km/h.

2.3. Simuladores de Direção Como Ferramenta de Pesquisa

Simuladores de direção tem sido utilizados como ferramenta de estudos com os mais variados objetivos nos últimos anos. Sua validade tem sido constatada com frequência, através de estudos comparativos entre dados coletados no mundo real e em ambiente simulado (Klee et al., 1999; Bella, 2008; Hussain et al. 2019).

As plataformas de simuladores diferenciam-se entre imersivas e de tela plana, ou seja, sem a presença de óculos de realidade virtual. A plataforma imersiva permite ao usuário ter uma experiência com 360 graus de visão, apenas vendo e ouvindo o que se passa na simulação, portanto, promove uma maior sensação de estar no ambiente simulado (Mourant e Schultheis, 2001).

Essa ferramenta pode ser definida como versátil, dada a variedade de assuntos e diferentes abordagens com as quais é utilizada. Como análise de percepção da sinalização vertical (Larocca et al, 2018), os efeitos na condução enquanto se realiza outras atividades, como conversar, beber e fumar (Yannis et al. 2011), efeitos de cansaço por monotonia em estradas (Thiffault e Bergeron, 2003), efeitos da geometria de vias no comportamento do usuário (Wang e Wang, 2018) e influência de diferentes quantidades de álcool sobre indivíduos de faixas etárias distintas (Yadav e Velaga, 2019).

Pesquisas anteriores constataram que seu uso pode provocar sintomas de mal-estar em condutores durante o período de utilização, o chamado *Simulator Sickness* (Kennedy et al., 1993; Brooks et al., 2010; Classen et al., 2011; Torres et al., 2018), sendo essa a principal limitação do simulador. Nodari et al. (2017) constataram que a utilização por um período de 8 minutos produziu sintomas de mal-estar variando entre leves e moderados.

3. MÉTODO

Esta seção explicará os critérios utilizados no estudo. Será feita uma introdução das características do simulador de direção imersivo e do cenário utilizados. Após, será discutido o projeto de experimentos desenvolvido, com a devida apresentação dos dados usados neste estudo, também serão discutidos os métodos utilizados para a análise dos *clusters* e os demais procedimentos adotados em toda a realização do experimento.

3.1. Simulador de Direção

Os dados que compõe este estudo são provenientes de um experimento realizado em um simulador de direção do tipo imersivo. Os participantes voluntários utilizaram óculos de realidade virtual e fones de ouvido circo auricular para serem isolados do ambiente ao seu redor, tanto visualmente quanto acusticamente, e para que houvesse uma sensação de imersão causada pelo cenário apresentado. No ambiente virtual imersivo a condução se torna mais realista, pois os óculos de realidade virtual permitem ao usuário a possibilidade de enxergar em qualquer direção do ambiente rodoviário virtual, facilitando o processo de imersão. Neste experimento, o espaço para o motorista (*cockpit*) do simulador foi montado em uma plataforma dinâmica com dois graus de liberdade capaz de reproduzir as forças que atuam em um veículo em movimento (Figura 1).



Figura 1: Simulador de condução imersivo

Optou-se por configurar câmbio automático tanto para facilitar quanto uniformizar a condução entre os participantes. Os componentes do veículo (volante, pedal do freio, pedal do acelerador, espelhos e painel) são semelhantes àqueles encontrados em veículos de passeio.

3.2. Cenário de Estudo

O cenário apresentado foi originalmente utilizado no experimento desenvolvido por Oliveira e Nodari (2018), tratando-se de um trecho rodoviário dividido em dois segmentos, resultando em uma extensão total de aproximadamente 5700 metros. O primeiro segmento conta com uma extensão aproximada de 2500 metros de pista dupla em reta, e teve como objetivo adaptar o condutor ao simulador. O segundo segmento foi onde ocorreu o

desenvolvimento do experimento, contando com um percurso em relevo ondulado de pista simples, com aproximadamente 3200 metros de extensão de trechos retilíneos e curvos em sucessão. As larguras da faixa e do acostamento eram de 4 e 3 metros respectivamente, e não havia presença de defensas em nenhuma parte do trecho. Havia tráfego no sentido contrário composto por veículos leves e pesados, no entanto não havia tráfego no sentido em que o participante conduzia, não havendo necessidade de efetuar manobras de ultrapassagem. O experimento conduzido utilizou o resultado de Nodari et al. (2017) como limite de tempo de utilização do simulador, em virtude da sensação de mal-estar provocada por períodos prolongados. Os pesquisadores em questão chegaram à conclusão que períodos de utilização de aproximadamente 8 minutos tem impacto baixo na sensação de mal-estar, o chamado *simulator sickness*. O período de utilização neste experimento foi de aproximadamente 4 minutos, portanto com impactos reduzidos no mal-estar dos participantes.

3.3. Dados da Análise

O presente estudo foi realizado utilizando a base de dados obtida no experimento realizado por Oliveira e Nodari (2018). É importante notar que esse trabalho utilizou um projeto de experimentos fatorial 2^k , que resultou em um tamanho de amostra de 32 participantes, enquanto neste estudo será utilizado o tamanho total da amostra coletada, com 73 voluntários aleatórios de gênero e idades variados, em um total de 42 homens e 31 mulheres com idade média de 30 anos.

Os dados utilizados foram organizados em dois lotes: o lote 1 refere-se aos dados advindos de questionários preenchidos pelos voluntários. Ao todo foram respondidas 32 questões organizadas em dois questionários: (i) *Questionário de perfil do condutor* (Tabela 1), que conta com 10 questões e (ii) *Questionário de experiência no simulador* (Tabela 2), que conta com 22 questões. No presente estudo não foram utilizadas as questões QC6 (Quais características descrevem melhor o tipo de condutor que você é?) pois as possíveis respostas (Inseguro / Cauteloso / Confiante / Agressivo) não são necessariamente independentes entre si, e QC8 (Usa óculos durante a condução), por haver alguns condutores que utilizam óculos durante sua condução normal e que não puderam utilizá-lo durante o experimento, fazendo com que a variável não pudesse ser avaliada de forma apropriada.

Tabela 1: Questionário de perfil do condutor

<i>Cód.</i>	<i>Questão</i>
QC1	Gênero
QC2	Idade
QC3	Com que frequência você se entretém com jogos de corrida ou carros? *
QC4	Já participou de outras simulações de condução ou realidade virtual? (sim ou não)
QC5	Você se considera um condutor: (experiente ou inexperiente)
QC7	Tempo de carteira de habilitação
QC9	Em vias urbanas você dirige com que frequência? **
QC10	Em vias rurais você dirige com que frequência? **

*Respostas variando entre (Diariamente / Algumas vezes por semana / Algumas vezes por mês / Nunca)

**Respostas variando entre (Todos os dias / Algumas vezes por semana / Algumas vezes por mês / Algumas vezes por ano / Não dirijo)

O questionário referente à experiência no simulador conta com questões que buscam compreender a interação do condutor com o equipamento e o seu nível de imersão, e sua aplicação ocorreu logo após a utilização do simulador por parte do voluntário.

As 22 questões foram tabuladas com respostas de 6 pesos diferentes, de forma que a resposta que refletisse a experiência mais negativa (experiência ruim) teria peso 0 e a mais positiva teria peso 5 (experiência boa). O lote 2 de dados usados neste estudo refere-se àqueles coletados durante a experiência de condução dos voluntários no percurso do simulador, em que ocorreu a coleta dos dados de velocidade em estacas com 20 metros de distância entre si, para que pudessem ser utilizados os dados de velocidade média e seu desvio padrão.

Tabela 2: Questionário de experiência no simulador

<i>Cód.</i>	<i>Questão</i>
QA1	Conseguiu controlar (volante, pedal, etc.) o veículo simulado?
QA2	Como foi sua percepção de sensação de velocidade no cenário simulado?
QA3	Em função das respostas do simulador, qual o nível de realidade sentida para suas ações?
QA4	Qual o nível de naturalidade com o qual você interagiu com o simulador?
QA5	Em que nível o aspecto visual conseguiu te envolver no ambiente?
QA6	Como foi o seu controle dos mecanismos de direção do simulador?
QA7	Qual o nível de convicção que você daria para os objetos presentes na simulação?
QA8	Quão idêntico é o ambiente simulado comparado com uma experiência real?
QA9	Você foi capaz de prever o que aconteceria na simulação logo após executar alguma ação?
QA10	Quanto você conseguiu explorar visualmente o ambiente para tomar decisões baseadas no seu entorno?
QA11	Quantos detalhes você foi capaz de identificar nos objetos ao observá-los?
QA12	Você percebeu diferenças no nível de detalhamento de objetos a diferentes distâncias?
QA13	Quão envolvido com o ambiente virtual você esteve?
QA14	Como classifica o tempo entre uma ação tomada e a resposta do simulador?
QA15	O quanto se sentiu adaptado ao simulador?
QA16	No fim da simulação, como você classificaria seu nível de aptidão para dirigir no simulador?
QA17	A qualidade do vídeo interferiu na realização de alguma tarefa?
QA18	Os mecanismos de direção do cockpit interferiram na realização de alguma tarefa?
QA19	Quanto você conseguiu se concentrar efetivamente na simulação, ao invés de desviar a atenção para os mecanismos de direção?
QA20	Em que nível o aspecto auditivo conseguiu te envolver no ambiente?
QA21	Quão bem você conseguiu identificar os sons?
QA22	Conseguiu identificar as origens dos sons?

3.4. Procedimento da Análise

Este item se destina a abordar os procedimentos de análise de dados realizados, através de uma explicação resumida da metodologia. O procedimento iniciou com a realização de uma Análise Fatorial Exploratória, onde foram extraídos fatores relevantes para a Análise dos *Clusters*.

3.4.1. Análise Fatorial Exploratória

Como cada questão foi tratada como uma variável, o número total de variáveis para análise tornou-se elevado, levando-se em conta o reduzido tamanho amostral. Assim, procurou-se uma redução do número de variáveis para a posterior Análise de *Clusters*. A Análise Fatorial Exploratória (AFE), permite que seja obtido um conjunto de **variáveis correlacionadas** e

agrupa essas variáveis em fatores, os quais seriam variáveis não-observadas. Os testes de Esfericidade de Bartlett, Kayser-Meyer-Olkin e o gráfico *Scree Plot*, detalhados a seguir, foram utilizados com objetivo de validar a utilização da AFE.

- **Teste de Esfericidade de Bartlett**

Para uma correta utilização da AFE, é necessário que haja relação entre as variáveis componentes, por isso optou-se por utilizar o teste de esfericidade de Bartlett (Bartlett, 1950), que testa a hipótese de que as variáveis **não sejam correlacionadas**.

O teste verifica a hipótese de a matriz das correlações ser a matriz identidade, o que indicaria que todos os coeficientes de correlação são iguais a zero, indicando que não existe relação entre as variáveis e resultando em uma situação onde seria rejeitado o uso de uma AFE, portanto, espera-se rejeitar esta hipótese.

- **Teste de Kayser-Meyer-Olkin**

O segundo teste apresentado é o de Kaiser-Meyer-Olkin (KMO) que indica através de um índice, o quanto a amostra é adequada para uma AFE (Hair et al., 2005). O KMO varia entre 0 e 1, e valores entre 0,8 e 0,9 são considerados ótimos, valores entre 0,7 e 0,8 são bons, valores entre 0,5 e 0,7 são aceitáveis e valores abaixo de 0,5 indicam que deveriam ser coletados mais dados ou que as variáveis incluídas deveriam ser repensadas (Hutcheson e Sofroniou, 1999).

- **Gráfico de Escarpa (Scree Plot)**

O *Scree Plot* (Cattle, 1966) consiste em um gráfico utilizado visualmente para analisar quantos componentes, ou fatores, explicam a maior parte da variabilidade dos dados, mostrando o número de componentes fatoriais extraídos em relação aos autovalores associados a esses componentes. Geralmente, esse número equivale ao número de fatores anterior ao ponto de inflexão da curva do gráfico, quando o gráfico começa a ficar em linha quer dizer que os outros fatores já não explicam tanto assim.

3.4.2. *Análise de Clusters*

Após a formação dos fatores na AFE, estes foram usados na formação de *clusters* (ou agrupamentos). Os agrupamentos foram formados a partir das respostas dos indivíduos nas questões formadoras de fatores.

A *Análise de Clusters* é uma técnica exploratória que permite agrupar objetos com base nas características que eles possuem, colocando-os em grupos homogêneos relativamente à uma ou mais características em comum (Hair et al, 2005), diferindo da AFE que está prioritariamente interessada em agrupar variáveis. Foi utilizado o Agrupamento Hierárquico de *Clusters*, no qual ocorrem passos sucessivos de agregação dos sujeitos considerando-os individualmente, isto é, as duas observações mais próximas que ainda não estão agrupadas são combinadas.

Foi executada Análise de *Clusters* do tipo hierárquica com o método Ward, com a distância euclidiana ao quadrado como medida de dessemelhança. Os dados foram padronizados em uma escala de -1 até 1. A distância euclidiana ao quadrado, consiste na soma dos quadrados das diferenças e é a distância recomendada por Hair et al. (2005) para o método de Ward.

Segundo (Hair et al., 2005), o método de Ward consiste em um procedimento de agrupamento hierárquico no qual a medida de similaridade usada para juntar agrupamentos é calculada como a soma de quadrados entre os dois agrupamentos feita sobre todas as variáveis. Esse método tende a resultar em agrupamentos de tamanhos aproximadamente iguais. Em cada estágio, combinam-se os dois agrupamentos que apresentarem menor aumento na soma global de quadrados dentro dos agrupamentos. Assim, utilizou-se o método Ward, que tem se revelado um dos melhores e mais usados métodos hierárquicos de aglomeração (Malhotra, 2006).

• **Teste Qui-Quadrado**

Para verificar se uma distribuição de perfis está associada à distribuição dos *clusters*, realizou-se o teste Qui-Quadrado de aderência, simbolizado por χ^2 . Esse é um teste de hipóteses que se destina a encontrar as possíveis divergências entre as frequências observadas e esperadas para um certo evento, e se as proporções observadas desses eventos mostram ou não diferenças significativas (Bolboacă, 2011). É um teste não paramétrico, ou seja, não depende dos parâmetros populacionais como média e variância e pode-se dizer que dois grupos se comportam de forma semelhante se as diferenças entre as frequências observadas e as esperadas em cada categoria forem próximas a zero.

O pesquisador trabalha com uma hipótese nula, que afirma que as frequências observadas não são diferentes das frequências esperadas e com a hipótese alternativa que afirma o contrário. Adotou-se o valor- $p < 0,05$ (adotando-se o nível de significância de 5%), fazendo com que nesses casos rejeite-se a hipótese nula (H_0) de independência entre as variáveis.

• **Teste ANOVA**

O teste ANOVA (Análise Univariada de Variância) tem como objetivo comparar a média de população amostral, e assim identificar se essas médias diferem significativamente entre elas e não por erro amostral (Hair et al., 2005). A diferença básica entre os testes de hipótese e a análise de variância é o número de amostras. Enquanto nos testes de hipótese se trabalha com duas amostras, a ANOVA compara a média de mais de duas amostras e determina se ao menos uma se difere significativamente das demais.

Dentro do teste ANOVA serão utilizados alguns testes estatísticos adicionais, como o de Levene, que verifica se os grupos possuem variância homogênea (requisito para o teste

ANOVA) e o teste DMS (Mínima Diferença Significativa) de comparações múltiplas, que aponta quais grupos tiveram médias significativamente diferentes.

4. RESULTADOS

Nesta sessão serão apresentados os resultados alcançados através dos métodos citados, com apresentação dos fatores gerados na AFE e análise dos parâmetros dos *clusters*.

4.1. Análise Fatorial Exploratória (AFE)

A AFE permitiu que o conjunto original de 32 variáveis (questões) fosse analisado e organizado em um número fatores, resultando em 5 fatores que englobam 12 das variáveis originais. Esses fatores podem ser entendidos como variáveis não observadas. A Tabela 3 apresenta os cinco fatores extraídos para o modelo final, a primeira coluna apresenta os nomes dos fatores definidos pelo pesquisador, escolhidos como uma forma de sintetizar as características das variáveis em um fator relevante para o estudo. A segunda coluna apresenta as variáveis que ficaram mais correlacionadas na AFE, também foram considerados os resultados dos métodos de KMO e Bartlett para definir o modelo final.

Tabela 3: Fatores Extraídos

<i>Imersão (Fator 1)</i>	QA3	<i>Prática (Fator 2)</i>	QC9	<i>Qualidade do Equipamento (Fator 3)</i>	QA17	<i>Interação com a Realidade Virtual (Fator 4)</i>	QA4	<i>Envolvimento com a Simulação (Fator 5)</i>	QA13
	QA5		QA18		QA10		QA19		
	QA7								
	QA8								

• Testes de Kayser-Mayer-Olkin e Bartlett

O valor de significância (valor-p) mais utilizado para rejeitar a hipótese nula é de 5%, o que significa que um resultado menor que 0,05 pode ser considerado significativo, rejeitando-se a hipótese nula. O teste de esfericidade de Bartlett obteve um valor-p abaixo de 10^{-4} e o teste KMO obteve escore de 0,584, portanto considerado ajuste aceitável. A variância total explicada observada indica que os 5 fatores extraídos explicam 70,92% dos dados originais.

• Gráfico de Escarpa (Scree Plot)

Como pode ser observado Figura 2, o número de componentes anterior ao ponto de inflexão é de cinco, desta forma, o gráfico indica a extração de cinco fatores, confirmando o número de 5 fatores retidos nesta AFE.

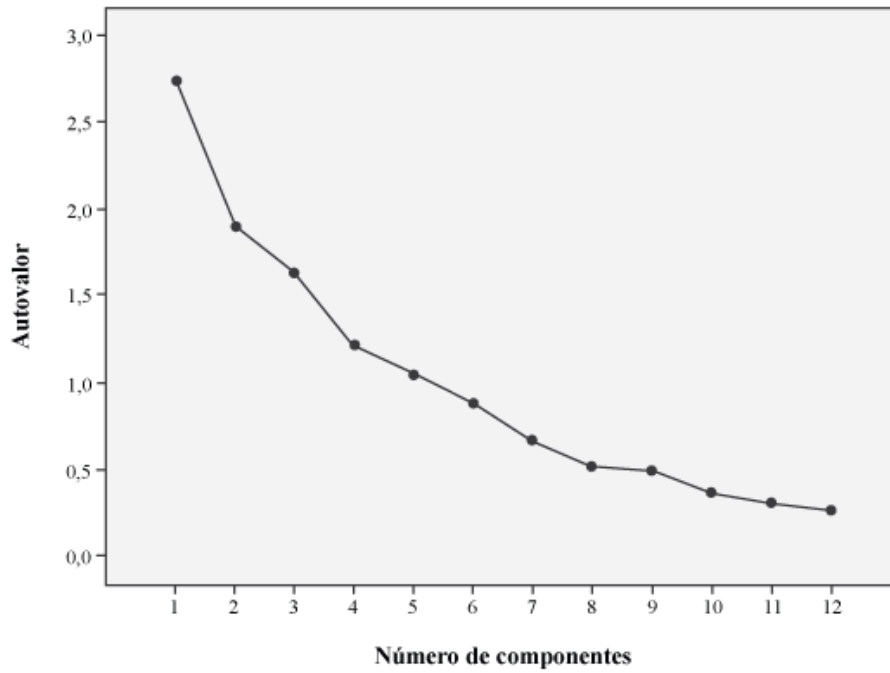


Figura 2: Gráfico de Escarpa

4.2. Análise de Clusters

A partir dos dados dos respondentes em relação às variáveis dos fatores extraídos, foram gerados *clusters* a partir do método de Ward. Observando-se o dendograma (Figura 3) que também é uma forma de representar o processo de formação de *clusters*, visualizando os retângulos maiores (com distância euclidiana quadrada inferior a 15) que unem as ramificações, podem ser identificados 3 agrupamentos com este método. Com estas constatações, serão retidos 3 *clusters*, pelo método Ward.

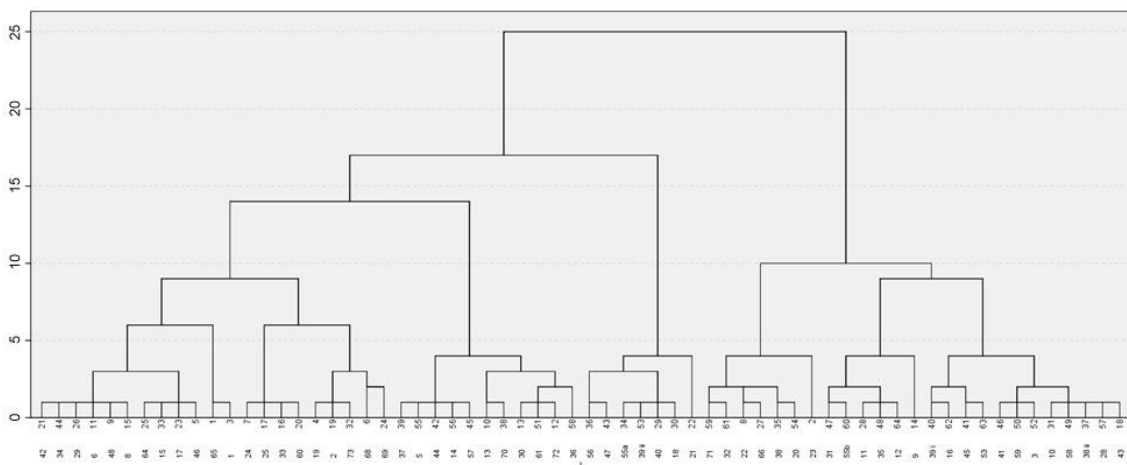


Figura 3: Dendograma de formação dos *clusters*

4.2.1. Perfis dos Clusters

A partir da criação dos *clusters* buscou-se uma identificação dos seus perfis, com relação à frequência. A Tabela 4 apresenta a distribuição dos indivíduos em cada *cluster*, 50% dos indivíduos ficaram no *cluster* 1, 39% no *cluster* 2 e 11% no *cluster* 3. A análise será feita de acordo com as características de perfil dos indivíduos que compõem cada *cluster* e com os dados de velocidade média.

Tabela 4: Distribuição de indivíduos por *cluster*

<i>Cluster</i>	<i>Nº de indivíduos</i>	<i>Porcentagem</i>
1	32	50%
2	25	39%
3	7	11%
Total	64	100,0

- *Clusters por Sexo*

Analisando-se os *clusters* por sexo (Tabela 5), pode-se observar que o *cluster* 2 possui mais mulheres que homens (52% é do sexo feminino), mas de qualquer forma não há grande diferença entre a distribuição dos homens e mulheres nos *clusters* 1 e 2, já os integrantes do *cluster* 3 são na sua maioria do sexo masculino, 5 homens e duas mulheres, ou seja, 71% dos integrantes do sexo masculino.

Tabela 5: *Clusters* por sexo

<i>Clusters</i>			<i>Sexo</i>		
			<i>Feminino</i>	<i>Masculino</i>	<i>Total</i>
<i>Cluster 1</i>	Contagem		14	18	32
<i>Cluster 2</i>	Contagem		13	12	25
<i>Cluster 3</i>	Contagem		2	5	7
	<i>Total</i>		29	35	64

Para verificar se a distribuição dos sexos está estatisticamente associada à distribuição dos *clusters*, realizou-se o teste Qui-Quadrado de independência, como valor-p do teste foi 0,550, o teste não foi significativo, então pode-se concluir que não houve associação significativa entre os grupos. Isso significa dizer que os *clusters* não podem ser considerados estatisticamente diferentes em relação ao sexo de seus integrantes.

Mesmo não havendo significância no teste, descritivamente essa amostra mostrou que o único grupo com mais mulheres presentes foi o *cluster 2*, e no *cluster 3* ocorreu a maior diferença constatada, sendo composto em sua maioria por homens.

- *Clusters por Idade*

A seguir, serão analisados os *clusters* pela média aritmética da idade dos seus integrantes. Na Tabela 6 pode-se observar que a média de idade é diferente nos três *clusters*, sendo o *cluster 1* o mais jovem em média, o *cluster 2* teve uma média intermediária e o 3 obteve a média de idade mais alta.

Tabela 6: *Clusters por idade*

	<i>Nº de indivíduos</i>	<i>Média (anos)</i>	<i>Desvio Padrão</i>	<i>Mínimo</i>	<i>Máximo</i>
<i>Cluster 1</i>	32	26,688	10,1677	20	65
<i>Cluster 2</i>	25	32,760	10,2643	21	59
<i>Cluster 3</i>	7	35,571	10,2446	24	52
Total	64	30,031	10,6323	20	65

Primeiramente foi executado o Teste Levene para verificar a suposição de homogeneidade de variâncias, pressuposto para a execução do teste ANOVA. Como o valor-p do teste foi 0,661, que é maior que 0,05, não se rejeita esta hipótese. Dessa forma, o teste ANOVA pode ser utilizado. O teste ANOVA teve um valor-p = 0,032, demonstrando que o teste foi significativo, o que indica que há diferenças estatisticamente significativas entre as médias de idade dos *clusters*.

Para verificar quais *clusters* se diferenciaram entre si com relação à idade média, realizou-se o teste de comparações múltiplas pelo método de mínima diferença significativa (DMS) que demonstrou a existência de diferenças significativas entre o *cluster 1* e os demais.

- *Clusters por Tempo de Carteira de Habilitação (QC7)*

A Tabela 7 apresenta os dados dos *clusters* para o tempo de carteira de habilitação:

Tabela 7: *Clusters por tempo carteira de habilitação*

	<i>Nº de indivíduos</i>	<i>Média (anos)</i>	<i>Desvio Padrão</i>	<i>Mínimo</i>	<i>Máximo</i>
<i>Cluster 1</i>	32	7,50	9,850	1	47
<i>Cluster 2</i>	25	13,80	10,231	2	39
<i>Cluster 3</i>	7	17,29	10,484	6	34
Total	64	11,03	10,574	1	47

Como pode-se observar, as médias do tempo de carteira de habilitação são diferentes para os três *clusters*. O teste Levene não rejeitou a hipótese de igualdade de variâncias (valor-p=0,429), e o teste ANOVA obteve valor-p=0,018, portanto foi significativo, indicando que existem diferenças estatisticamente significativas das médias da variável “tempo de carteira de habilitação” entre os *clusters*. Com a utilização do teste DMS, pode-se concluir que o *cluster* 1 se diferenciou estatisticamente dos outros, enquanto os *clusters* 2 e 3 não se diferenciaram. O *cluster* 1 é formado pelos indivíduos menos experientes, em média, o resultado indicou um tempo de habilitação significativamente diferente dos demais *clusters*, apesar dos *clusters* 2 e 3 terem médias diferentes, essa diferença não foi considerada significativa do ponto de vista estatístico.

- *Clusters para Você se Considera um Condutor Experiente ou Inexperiente? (QC 5)*

A Tabela 8 apresenta os dados de análise para a variável QC5 – Você se considera um Condutor Experiente ou Inexperiente. Conforme pode-se observar, condutores auto-identificados como inexperientes só foram observados no *cluster* 1, onde aproximadamente 40% se identificaram como inexperientes.

O teste Qui-Quadrado foi altamente significativo (valor-p<0,01), indicando que há associação entre as variáveis, seus resíduos padronizados indicaram que há mais condutores inexperientes no *cluster* 1 do que esperado.

Tabela 8: *Clusters por “Você se Considera um Condutor Experiente ou Inexperiente”*

		<i>Você se considera um condutor</i>			
		<i>Experiente</i>	<i>Inexperiente</i>	<i>Total</i>	
<i>Clusters</i>	<i>Cluster 1</i>	Contagem	19	13	32
		%	59,4%	40,6%	100,0%
	<i>Cluster 2</i>	Contagem	25	0	25
		%	100,0%	0,0%	100,0%
	<i>Cluster 3</i>	Contagem	7	0	7
		%	100,0%	0,0%	100,0%
<i>Total</i>		Contagem	51	13	64
		%	79,7%	20,3%	100,0%

- *Clusters por Velocidade Média*

A Tabela 9 apresenta os dados dos 3 *clusters* para a variável Velocidade Média. Nesse sentido, é importante notar que o número de condutores pertencentes à agrupamentos analisados para essa variável resultou num total de 56, devido a 8 condutores que foram retirados da análise por se envolverem em colisões durante o percurso simulado, o que comprometeu seus dados de

velocidade média.

Tabela 9: *Clusters* por Velocidade Média

	<i>Nº de indivíduos</i>	<i>Média (Km/h)</i>	<i>Desvio Padrão</i>	<i>Mínimo</i>	<i>Máximo</i>
<i>Cluster 1</i>	28	79,4123	13,7952	61,35	117,17
<i>Cluster 2</i>	21	83,4941	15,2622	56,62	115,34
<i>Cluster 3</i>	7	81,7999	9,5347	67,01	93,49
Total	56	81,2414	13,846	56,62	117,17

O teste Levene não rejeitou a hipótese de igualdade de variâncias (valor-p=0,440), porém o teste ANOVA apresentou valor-p = 0,598, indicando que o teste não foi significativo, e que não existem diferenças estatisticamente relevantes das médias da variável “velocidade média” entre os *clusters*. Apesar de as médias de velocidade entre os três *clusters* apresentarem valores muito semelhantes, é importante notar que em virtude do desvio padrão mais elevado, o *cluster 2* apresentou a situação mais crítica, enquanto o *cluster 3* apresentou a situação de menor criticidade e o *cluster 1* apresentou dados intermediários. Como demonstrado por Oh e Ritchie (2005), o desvio padrão é um bom indicador de condições de tráfego perigosas, e como indicado por Oh et al. (2005) e Ahmed et al. (2012), um maior desvio padrão de velocidades é um indicativo de uma maior incidência de acidentes. Apesar das diferenças nas médias, do ponto de vista estatístico elas não foram significativas.

4.3. Análise dos *Clusters* pelos Fatores Extraídos

Neste tópico será apresentada a análise dos *clusters* segundo os fatores gerados na AFE. A Figura 4 apresenta o gráfico com as médias dos *clusters* para cada fator em uma escala padronizada de -1 até 1. Pode-se observar, por exemplo, que o *cluster 2* ficou com média superior em todos os fatores, exceto “interação com a realidade virtual”, onde o *cluster 3* se destacou. O *cluster 1* ficou com média bem abaixo no fator “prática”. A menor média observada foi a do *cluster 3* no fator envolvimento.

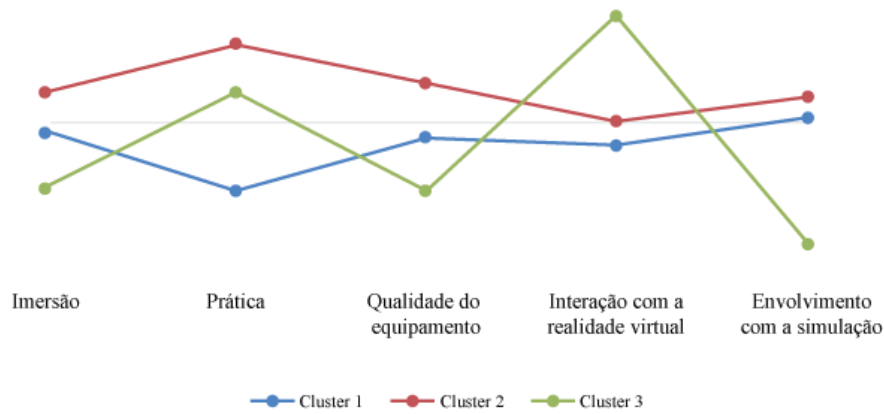


Figura 4: Médias dos *clusters* para os 5 Fatores

Os próximos itens apresentarão as análises de cada fator individualmente.

4.3.1. Fator Imersão

O teste Levene não rejeitou a hipótese de igualdade de variâncias (valor- $p=0,691$), portanto pode-se proceder ao teste ANOVA. Com valor- $p = 0,035$, o teste ANOVA obteve um resultado significativo, indicando que existem diferenças estatisticamente significativas das médias dos *clusters* para o fator “imersão”. O teste DMS verificou que os *clusters* 2 e 3 se diferenciaram significativamente ($p=0,015$). A Figura 5 apresenta a pontuação média de cada *cluster* para as variáveis do fator 1:

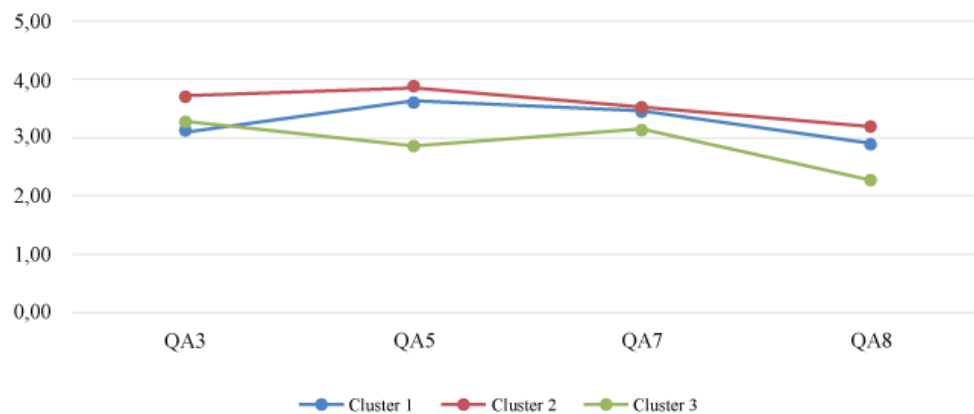


Figura 5: Pontuação média para o Fator 1 - "Imersão"

Como pode-se notar, as médias foram mais altas para o *cluster* 2 em todas as questões, indicando que este grupo teve uma impressão melhor do simulador que os demais quando comparado a um cenário real, os usuários deste grupo indicaram um maior nível de convicção com relação aos objetos da simulação e um maior envolvimento visual. O *cluster* 1 apresentou

uma experiência parecida com o 2, porém com escore mais baixo na variável QA3, indicando que os membros desse grupo sentiram menos realidade em relação às respostas do simulador. O *cluster* 3 obteve médias mais baixas em praticamente todas estas questões.

4.3.2. Fator Prática

O teste Levene apresentou valor-p = 0,156, portanto não rejeitou a hipótese de igualdade de variâncias. O resultado do teste ANOVA, para verificar se as diferenças são significativas entre os *clusters* no Fator 2 teve um valor muito baixo, inferior a 10^{-4} , indicando teste significativo.

De acordo com o teste DMS, o *cluster* 1 apresentou diferenças significativas em relação aos *clusters* 2 e 3 ($p \leq 10^{-4}$) e os *clusters* 2 e 3 não se diferenciaram entre si. A Figura 6 apresenta a pontuação média dos *clusters* para o Fator 2:

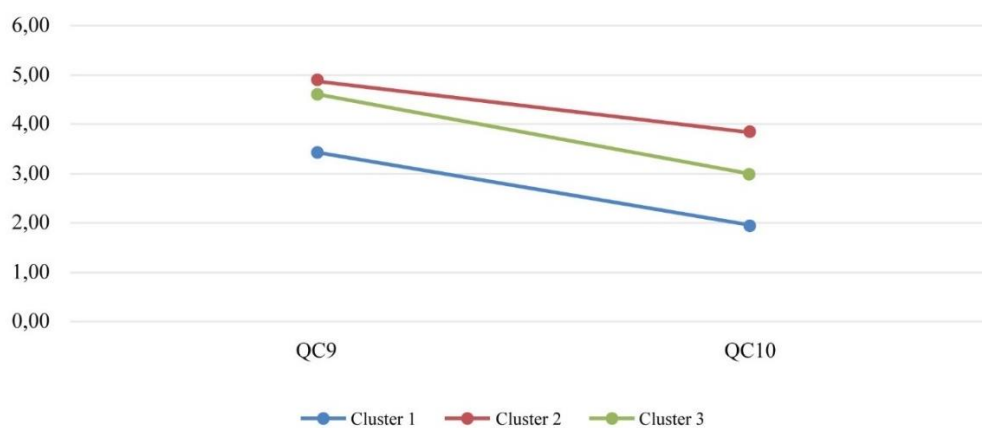


Figura 6: Pontuação média para o Fator 2 - "Prática"

Maiores notas, em média, para o *cluster* 2 em todas as questões, demonstrando que este *cluster* dirige com mais frequência que os demais tanto em vias urbanas quanto em vias rurais. O *cluster* 1 obteve o resultado oposto, mostrando médias mais baixas em ambas questões, o que demonstra que este agrupamento tem um menor costume de dirigir que os demais. O *cluster* 3 teve médias intermediárias.

É interessante destacar que os três grupos obtiveram um escore mais alto na variável "QC9 – Em vias Urbanas você dirige com que frequência?" do que na variável QC10, que perguntava sobre vias rurais, e que o *cluster* 1 teve resultados bastante baixos levando em conta as respostas disponíveis no questionário, onde, por exemplo, um escore de 2 equivale a dirigir algumas vezes por ano nas vias em questão, e um escore igual a 3 equivale a dirigir algumas vezes por mês.

4.3.3. Fator Qualidade do Equipamento

O teste Levene não rejeitou a hipótese de igualdade de variâncias (valor-p=0,095), portanto pode-se proceder ao teste ANOVA. Com valor-p=0,008, o teste ANOVA foi significativo, indicando que existem diferenças estatisticamente significativas das médias dos *clusters* para o Fator “qualidade do equipamento”. O teste DMS indicou a existência de diferenças significativas entre os *clusters* 1 e 2 (p=0,022) e entre os *clusters* 2 e 3 (p=0,05). A Figura 7 indica a pontuação média dos *clusters* para as variáveis do Fator 3:

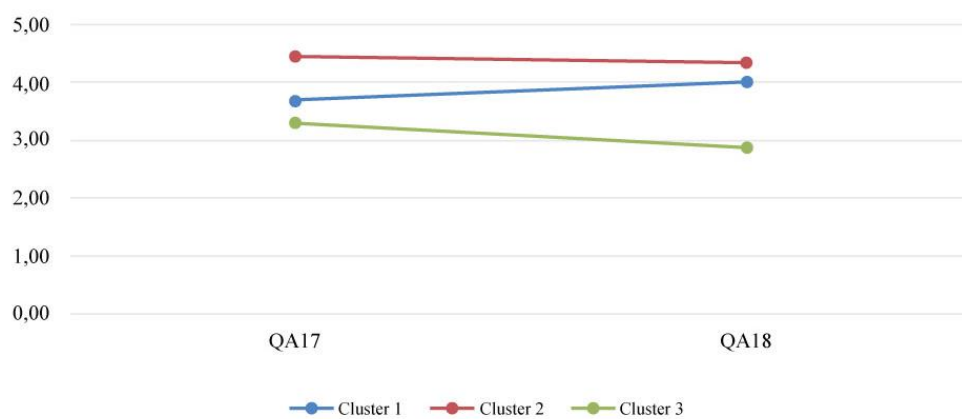


Figura 7: Pontuação média para o Fator 3 - "Qualidade do Equipamento"

Maiores notas em todas as questões, em média, para *cluster* 2, indicando que os indivíduos desse grupo tiveram uma melhor experiência com a qualidade de vídeo e que tiveram uma menor sensação que os mecanismos de direção atrapalharam em alguma tarefa, em relação aos demais *clusters*. O *cluster* 3 apresentou médias mais baixas, em todas as questões. O *cluster* 1 apresentou médias intermediárias.

4.3.4. Fator Interação com a Realidade Virtual

O teste Levene não rejeitou a hipótese de igualdade de variância (valor-p = 0,184), portanto pode-se proceder para o teste ANOVA. Com valor-p=0,003, o teste ANOVA foi significativo, indicando que existem diferenças estatisticamente significativas das médias dos *clusters* para o Fator “interação com a realidade virtual”. O teste DMS verificou a existência de diferenças significativas entre o *cluster* 3 e os demais. A Figura 8 apresenta o gráfico com a pontuação média dos *clusters* para as variáveis do Fator 4.

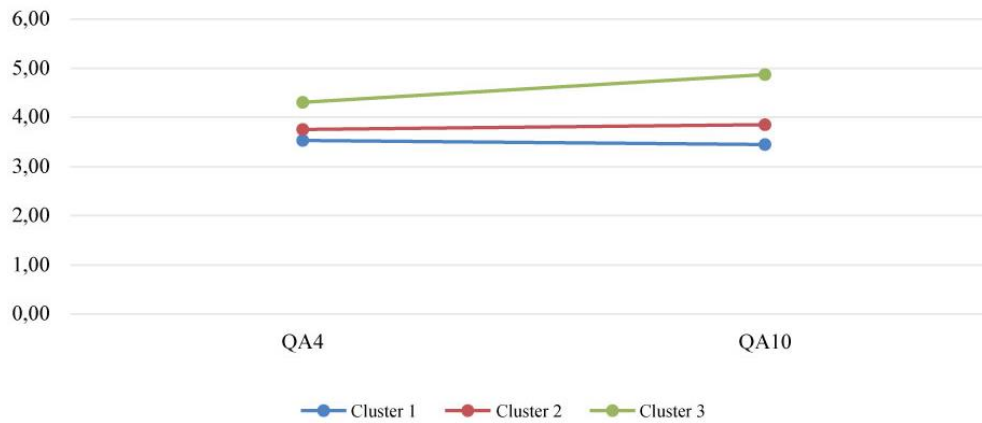


Figura 8: Pontuação média para o Fator 4 - "Interação com a Realidade Virtual"

O *cluster 3* apresentou maiores médias em ambas as variáveis, indicando que os participantes desse grupo sentiram uma melhor capacidade de explorar o ambiente visualmente e interagir naturalmente com o simulador. O *cluster 1* apresentou médias mais baixas, porém os resultados mostram uma experiência similar à do *cluster 2*.

4.3.5. Fator Envolvimento com a Simulação

O teste Levene não rejeitou a hipótese de igualdade de variâncias (valor-p = 0,390), pode-se proceder para o teste ANOVA. Com valor-p $\leq 10^{-4}$, o teste ANOVA foi significativo, indicando que existem diferenças estatisticamente significativas das médias dos *clusters* para o Fator "envolvimento com a simulação". O teste DMS indicou que o *cluster 3* se diferenciou significativamente dos demais com relação à média do Fator "envolvimento com a simulação". A Figura 9 apresenta as médias dos três *clusters* para o Fator 5.

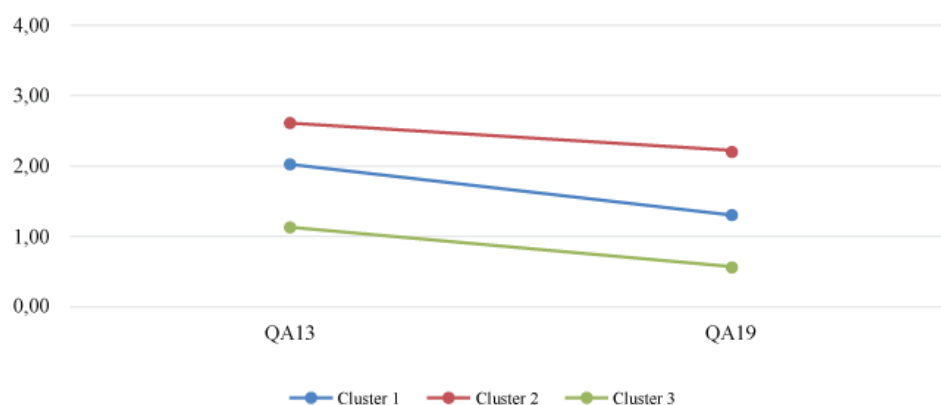


Figura 9: Pontuação média para o Fator 5 - "Envolvimento com a Simulação"

O *cluster 2* apresentou médias mais altas para ambas as variáveis do Fator 5, indicando que os membros deste *cluster* se sentiram menos distraídos pelos mecanismos de direção e

tiveram um melhor envolvimento com o mundo virtual no geral que os demais. O *cluster 1* apresentou médias intermediárias e o *cluster 3* apresentou médias mais baixas, o que indica uma experiência mais negativa com relação aos mecanismos de direção e um menor envolvimento com a simulação, o que coincide com os resultados do Fator “imersão”.

Na Tabela 10 pode-se observar os *clusters* extraídos e suas características:

Tabela 10: Quadro síntese dos *Clusters* e suas Características

<i>Cluster</i>	<i>Características Marcantes</i>	<i>Fatores de destaque</i>
1	Média de idade significativamente mais baixa	Média significativamente mais baixa no Fator prática
	Tempo de carteira de habilitação significativamente mais baixo	Média mais baixa no Fator interação com a realidade virtual
	40% dos condutores se consideram inexperientes	Médias intermediárias nos demais fatores
2	Média de idade de 33 anos	Médias mais altas em todos os fatores, exceto interação com a realidade virtual
	100% dos condutores se consideram experientes	Média significativamente mais alta que o <i>cluster 3</i> no Fator imersão
	Desvio padrão da velocidade mais alto entre os 3 <i>clusters</i>	Média significativamente mais alta no Fator qualidade do equipamento
3	71,4% dos condutores do sexo masculino	Médias mais baixas nos fatores imersão, qualidade do equipamento e envolvimento com a simulação
	Maior média de idade entre os <i>clusters</i>	
	100% dos condutores se consideram experientes	Média significativamente mais baixa que os demais no fator envolvimento com a simulação
	Menor desvio padrão da velocidade média entre os 3 <i>clusters</i>	Média significativamente mais alta no fator interação com a realidade virtual

5. CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo servir como um identificador de perfis de condutores através de suas características autorrelatadas, através de um experimento em ambiente simulado e das respostas dos usuários para dois tipos de questionários diferentes, incluindo questões sobre as características dos voluntários e da sua experiência com o simulador, procurando também auxiliar na melhoria desta ferramenta no futuro.

A análise resultou em três diferentes agrupamentos (*clusters*) de condutores, onde destacam-se algumas características:

- O *cluster* 1 foi formado por condutores com média de idade significativamente menor que os demais, cerca de 27 anos, e com tempo de carteira de habilitação significativamente menor, além de serem o único grupo onde usuários se identificaram como inexperientes, cerca de 40%. Os condutores deste *cluster* informaram que dirigem, em média, com menos frequência que os demais, tanto em vias urbanas quanto em vias rurais, estas diferenças foram significativas de acordo com o teste de DMS. O *cluster* também apresentou uma experiência mais negativa com relação às respostas do simulador em relação aos demais, apesar de a experiência no fator “imersão” ser similar ao *cluster* 2 nas demais variáveis.
- O *cluster* 2 foi formado por condutores de idade média intermediária entre os três *clusters*, cerca de 32 anos de idade. Este *cluster* apresentou o maior desvio padrão de velocidade e informou uma maior frequência ao volante que os demais, tanto em vias urbanas quanto em vias rurais. Constatou-se uma média significativamente maior no fator qualidade do equipamento, indicando uma melhor impressão da qualidade de vídeo e menor impressão que os mecanismos de direção atrapalharam em alguma tarefa. O *cluster* teve o escore mais alto em 8 das 10 questões voltadas para a experiência com o simulador, indicando que este *cluster* teve a experiência mais positiva.
- O *cluster* 3 foi formado por 7 condutores, sendo o menor dos 3 *clusters*. Este *cluster* foi o único que apresentou um predomínio de um sexo sobre o outro, com cinco condutores homens e duas mulheres. Obteve-se uma média de idade maior que os demais e um desvio padrão da velocidade mais baixo. O teste DMS verificou significância no fator envolvimento com a simulação, onde o *cluster* 3 obteve a média mais baixa indicando uma experiência mais negativa com relação aos mecanismos de direção e um menor envolvimento com a simulação. O *cluster* 3 também apresentou a pior experiência, em média, nos fatores imersão e qualidade do equipamento, diferenciando-se significativamente do *cluster* 2 de acordo com o teste DMS.

É importante notar que o fato de o *cluster* apresentar poucos indivíduos e alguns fatores serem formados por poucas variáveis, são indicativos de resultados mais sensíveis, onde poucas

mudanças nas respostas do questionário poderiam originar resultados diferentes.

Alguns padrões podem ser constatados a partir dos resultados encontrados, como um menor costume de dirigir no dia a dia por parte de indivíduos mais jovens, tanto nas vias rurais quanto nas urbanas. Um grande grupo se identificou como “inexperiente”, apesar de a idade média de 27 anos e o tempo médio de carteira de habilitação de mais de 7 anos do *cluster 1* demonstrarem que não se trata de indivíduos que não tiveram tempo suficiente de se ambientarem ou de adquirirem o costume de dirigir, o mesmo comportamento não foi demonstrado nos outros agrupamentos.

O agrupamento composto por indivíduos maduros e com maioria do sexo masculino apresentou os resultados que indicam a opinião mais negativa quanto à utilização do simulador, os resultados obtidos a partir dos dados deste grupo indicam uma menor imersão, devido às respostas negativas quanto aos fatores “envolvimento”, “qualidade do equipamento” e “imersão”. Constatações como esta podem indicar que o uso de simuladores de condução para análises de segurança viária podem ter resultados menos confiáveis para esse perfil de integrantes, visto que não se sentem tão envolvidos pelo ambiente imersivo. No entanto, o tamanho da amostra pode estar gerando distorções, sendo necessárias novas análises.

REFERÊNCIAS

- Ahmed, M. M., Abdel-Aty, M., & Yu, R. (2012). Bayesian Updating Approach for Real-Time Safety Evaluation with Automatic Vehicle Identification Data. *Transportation Research Record: Journal of the Transportation Research Board*, 2280(1), 60–67
- Bartlett, M. S. (1950) Tests of significance in factor analysis. *British Journal of Psychology*, 1950, 3, 77- 88.
- Begg, D., & Langley, J. (2001). Changes in risky driving behavior from age 21 to 26 years. *Journal of Safety Research*, 32(4), 491–499.
- Bella, F. (2008). Driving simulator for speed research on two-lane rural roads. *Accident Analysis & Prevention*, 40(3), 1078–1087
- Bolboacă, S. D., Jäntschi, L., Sestraş, A. F., Sestraş, R. E., & Pamfil, D. C. (2011). Pearson-Fisher Chi-Square Statistic Revisited. *Information*, 2(3), 528–545.
- Brooks, J. O., Goodenough, R. R., Crisler, M. C., Klein, N. D., Alley, R. L., Koon, Wills, R. F. (2010). Simulator sickness during driving simulation studies. *Accident Analysis & Prevention*, 42(3), 788-796
- Caleffi, F., Lucchesi, S. T., Anzanello, M. J., Cybis, H. B. B. (2016). Influência das condições climáticas e de acidentes na caracterização do comportamento do tráfego em rodovias, *TRANSPORTES*, v. 24, n.4, p. 57-63
- Cattell, R. B. (1966). The Scree Plot Test for the Number of Factors. *Multivariate Behavioral Research*, 1, 140-161.
- Carvalho, C.H.R (2016) Mortes por acidentes de transporte terrestre no brasil: análise dos sistemas de informação do ministério da saúde acessado em http://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_a2212.pdf
- Classen, S.; M. Bewernitz e O. Shechtman (2011) Driving Simulator Sickness: An Evidence-Based Review of the Literature. *American Journal of Occupational Therapy*, v. 65, p. 33-179-188.
- Dahlen, E. R., & Ragan, K. M. (2004). Validation of the Propensity for Angry Driving Scale. *Journal of Safety Research*, 35(5), 557–563.
- DePasquale, J. P., Geller, E. S., Clarke, S. W., & Littleton, L. C. (2001). Measuring road rage: development of the Propensity for Angry Driving Scale. *Journal of Safety Research*, 32(1), 1–16.

- Harré, N., Brandt, T., & Dawe, M. (2000). The Development of Risky Driving in Adolescence. *Journal of Safety Research*, 31(4), 185–194.
- Hair, J. F., Tatham, R. L., Anderson, R. E., Black, W. Análise multivariada de dados. São Paulo: Bookman. 2005 (5ª ed).
- Hussain, Q., Alhajjaseen, W. K. M., Pirdavani, A., Reinoltsmann, N., Brijs, K., & Brijs, T. (2019). Speed perception and actual speed in a driving simulator and real-world: A validation study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 62, 637–650.
- Hutcheson, G. D. & Sofroniou, N. (1999). The multivariate social scientist: Introductory statistics using generalized linear models. London: Sage Publications.
- Kaçan, B., Findık, G., Üzümcüoğlu, Y., Azık, D., Solmazer, G., Ersan, Ö., Özkan, T., Lajunen, T., Öz, B., Pashkevici, A, G (2019). Driver profiles based on values and traffic safety climate and their relationships with driver behaviors. *Transportation Research Part F: Traffic Psychology and Behaviour*, 64, 246–259.
- Kennedy, R. S.; N. E. Lane; K. S. Berbaum e M. G. Lilienthal (1993) Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, v.3(3), p. 203–220.
- Klee, H., Bauer, C., Radwan, E., & Al-Deek, H. (1999). Preliminary Validation of Driving Simulator Based on Forward Speed. *Transportation Research Record: Journal of the Transportation Research Board*, 1689, 33–39.
- Larocca, A. P. C., Ribeiro, R. L., da Cruz Figueira, A., de Oliveira, P. T. M. e S., Lulio, L. C., & Rangel, M. A. C. (2018). Analysis of perception of vertical signaling of highways by drivers in a simulated driving environment. *Transportation Research Part F: Traffic Psychology and Behaviour*, 58, 471–487.
- Lucidi, F., Giannini, A. M., Sgalla, R., Mallia, L., Devoto, A., & Reichmann, S. (2010). Young novice driver subtypes: Relationship to driving violations, errors and lapses. *Accident Analysis & Prevention*, 42(6), 1689–1696.
- McCartt, A. T., Shabanova, V. I., & Leaf, W. A. (2003). Driving experience, crashes and traffic citations of teenage beginning drivers. *Accident Analysis & Prevention*, 35(3), 311–320.
- Massie, D. L., Campbell, K. L., & Williams, A. F. (1995). Traffic Accident involvement rates by driver age and gender. *Accident Analysis & Prevention*, 27(1), 73–87.
- Malhotra, N. Pesquisa de marketing: uma orientação aplicada. Trad. Laura Bocco. 4 ed. Porto Alegre: Bookman, 2006.
- Musicant, O., Laufer, I., & Botzer, A. (2019). Changes in physiological indices and deceleration behaviour as functions of braking demands and driver's physiological cluster. *Transportation Research Part F: Traffic Psychology and Behaviour*, 62, 406–415.
- Mourant, R. R. e M. T. Schultheis (2001) A HMD Based Virtual Reality Driving Simulator. *Proceedings of the First International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*. Aspen, USA.
- Navas, J. F., Martín-Pérez, C., Petrova, D., Verdejo-García, A., Cano, M., Sagripanti-Mazuquín, O., Perandrés-Gómez, A., López-Martín, A., Cordovilla-Guardia, S., Megías, A., Perales, J.C., Vilar-López, R. (2019). Sex differences in the association between impulsivity and driving under the influence of alcohol in young adults: The specific role of sensation seeking. *Accident Analysis & Prevention*, 124, 174–179.
- Nodari, C. T.; M. C. Oliveira; M. R. Veronez; F. Bordin; L. Gonzaga Jr.; A.P.C. Larocca; C. Framarim (2017) Avaliação do Realismo e da Sensação de Mal-Estar (Simulator Sickness) no uso de Simulador Imersivo de Direção. *Anais do XXXI Congresso de Pesquisa e Ensino em Transportes, ANPET, Recife, v.1, p. 3103-3115*.
- Oliveira, M. C.; C. T. Nodari (2018) Análise da Posição Lateral sob uma Perspectiva de Segurança Viária. In: Oliveira, M.C.; Efeitos da Luz do Ambiente e de Sinalização Horizontal na Segurança em Rodovias: Estudo em Simulador de Direção Imersivo. Porto Alegre. UFRGS.
- Oltedal, S., & Rundmo, T. (2007). Using cluster analysis to test the cultural theory of risk perception. *Transportation Research Part F: Traffic Psychology and Behaviour*, 10(3), 254–262.
- Oh, C., Ritchie, S. G., & Chang, M. (2005). Real-Time Estimation of Accident Likelihood for Safety Enhancement. *Journal of Transportation Engineering*, 131(5), 358–363.
- Oh, C., Oh, J.-S., & Ritchie, S. G. (2005). Real-Time Hazardous Traffic Condition Warning System: Framework and Evaluation. *IEEE Transactions on Intelligent Transportation Systems*, 6(3), 265–272.
- Özkan, T., & Lajunen, T. (2005). Why are there sex differences in risky driving? the relationship between sex and gender-role on aggressive driving, traffic offences, and accident involvement among young turkish drivers. *Aggressive Behavior*, 31(6), 547–558.
- Perchonok, K. (1972) Accident cause analysis. CAL Report. ZM-5010-V-3, Cornell Aeronautical Laboratory, Inc.
- Reason, J., Manstead, A., Stradling, S., Baxter, J., & Campbell, K. (1990). Errors and violations on the roads: a real distinction? *Ergonomics*, 33(10-11), 1315–1332.

- Retchin, S.M., Anapolle, J. (1993). An Overview of the Older Driver. *Clinics in geriatric medicine*. 9. 279-96.
- Schmid Mast, M., Sieverding, M., Esslen, M., Graber, K., & Jäncke, L. (2008). Masculinity causes speeding in young men. *Accident Analysis & Prevention*, 40(2), 840–842.
- Sullman, M. J. M., Paxion, J., & Stephens, A. N. (2017). Gender roles, sex and the expression of driving anger. *Accident Analysis & Prevention*, 106, 23–30.
- Thiffault, P., & Bergeron, J (2003). Monotony of road environment and driver fatigue: a simulator study. *Accident Analysis & Prevention*, 35(3), 381–391.
- Torres, T. B.; L. B. Kappler; C. T. Nodari e C. Framarim (2018) O Mal-estar em Condução Simulada: Comparação Entre Simuladores Imersivos de Direção com Plataforma Estática e Dinâmica. *Anais do XXXII Congresso de Pesquisa e Ensino em Transportes*, ANPET, p. 3474-3486.
- Treat, J.R.; N.S. Tumbas; S.T. Mcdonald; D. Shinar; R.D. Hume; R.E. Mayer; R.L. Stansifer e N.J. Castellan (1979) Tri-Level Study Of The Causes Of Traffic Accidents Executive Summary
- Ulleberg, P. (2001). Personality subtypes of young drivers. Relationship to risk-taking preferences, accident involvement, and response to a traffic safety campaign. *Transportation Research Part F: Traffic Psychology and Behaviour*, 4(4), 279–297.
- World Health Organization, (2017). Gender, Women and Health: What Do We Mean by "sex" and "gender"? Acessado em 20 de junho de 2019 em: <https://www.who.int/gender-equity-rights/en/#>
- World Health Organization (2018) – Global Status Report on Road Safety 2018 Acessado em 10 de junho de 2019 em: https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/
- Wang, X., & Wang, X. (2018). Speed change behavior on combined horizontal and vertical curves: driving simulator-based analysis. *Accident Analysis & Prevention*, 119, 215–224.
- Yadav, A. K., & Velaga, N. R. (2019). Modelling the relationship between different Blood Alcohol Concentrations and reaction time of young and mature drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, 64, 227–245.
- Yannis, G.; E. Papadimitriou; C. Bairamis e V. Sklias (2011) Is it Risky to Talk, Eat or Smoke While Driving? Findings from a Driving Simulator Experiment, *Submitted to the 3rd International Conference on Road Safety and Simulation, September 14-16, 2011, Indianapolis, USA*
- Zuckerman, M. (1979). Sensation seeking: Beyond the optimal level of arousal.(1ª ed.). Psychology Press. London.