



Universidade de Brasília

Instituto de Psicologia

Programa de Pós-Graduação em Psicologia Social, do Trabalho e das Organizações

**Investigação de estilos de aprendizagem autodeclarados e observados: usando
psicometria, *big data* e *machine learning***

Mestrado

Rafael de Melo Balaniuk

Brasília, DF

2019



Universidade de Brasília

Instituto de Psicologia

Programa de Pós-Graduação em Psicologia Social, do Trabalho e das Organizações

**Investigação de estilos de aprendizagem autodeclarados e observados: usando
psicometria, *big data* e *machine learning***

Mestrado

Rafael de Melo Balaniuk

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Psicologia Social, do Trabalho e das Organizações como requisito parcial para obtenção do Título de Mestre em Psicologia.

Orientador: Prof. Dr. Jairo Eduardo Borges-Andrade

Brasília, DF, fevereiro de 2019

UNIVERSIDADE DE BRASÍLIA

Instituto de Psicologia

Essa dissertação foi aprovada pela seguinte Banca Examinadora:

Prof. Dr. Jairo Eduardo Borges-Andrade (Orientador)

Instituto de Psicologia – Universidade de Brasília

Prof. Dr. Josemberg Moura de Andrade (Membro Efetivo)

Instituto de Psicologia – Universidade de Brasília

Prof.^a Dr.^a Thaís Zerbini (Membro Efetivo)

Faculdade de Filosofia, Ciências e Letras – Universidade de São Paulo / Ribeirão Preto

Prof.^a Dr.^a Gardênia da Silva Abbad (Membro Suplente)

Instituto de Psicologia – Universidade de Brasília

*“Educação é a arma mais
poderosa que você pode usar
para mudar o mundo.”*

(Nelson Mandela)

AGRADECIMENTOS

Não foi nada fácil conciliar um trabalho em tempo integral com um mestrado acadêmico e tenho muitas pessoas a agradecer por me apoiarem nesse desafio. Primeiramente, agradeço a toda minha família, que é muito unida e sempre esteve comigo nos momentos tristes e felizes. Sei que nos últimos anos não fui tão presente, mas vocês não imaginam o quanto é bom ter vocês por perto. As viagens, os pontos Zé Mané e as tradições de família foram sempre incríveis.

Aos familiares do lado Balaniuk, agradeço pelo carinho que sempre me deram! Thelma, você não sabe como aquela ligação que você me fez onde, ao invés de brigar comigo por estar ausente em um momento importante, você perguntou como eu estava, mexeu comigo. Sua energia sempre me contagia e seu apoio me ajuda a continuar e a tentar ser uma pessoa melhor. Vó, te dar aquele abraço no fim do ano recarregou minhas energias como se eu nunca tivesse cansado! Seu carinho sempre aquece meu coração. Aninha, apesar de ser uma Medeiros, vejo nossas famílias como uma só. Obrigado não só por todos os momentos legais de conversas, almoços, churrascos... mas principalmente por distribuir amor nos momentos de brigas e ressentimentos. Obrigado por ajudar a reconciliar a família nos momentos difíceis! Caio, Bella e Lara, vocês não fazem ideia do quanto eu amo vocês! Desculpa por todos os domingos que eu não estive aí e pelos campeonatos de basquete e natação e apresentações de dança que eu perdi. Prometo que vou compensar!

Aos familiares do lado Melo, vocês são meu suporte nos momentos difíceis e nem tenho como agradecer o suficiente por isso! Bruno e Fred, obrigado por todas as semi-brigas de família que conseguimos nos esquivar fingindo que estávamos fazendo outra coisa. Mas obrigado também pela parceria nos momentos difíceis. Dri, agradeço muito por todo seu carinho nas fases difíceis. Apesar de não ser religioso, lembro muito de algumas rezas que você

puxou em momentos complicados e que serviram para acalmar e unir a família. Ana, você é um exemplo de profissional e de pessoa para mim! Obrigado por todas as conversas de sábado, mesmo sempre faltando comida na casa da vovó.

Vó, eu sei que onde você estiver, você está preocupada com o tanto que a gente tá comendo e se a gente levou ou não casaco. O que eu não tive chance de te dizer foi que, apesar de todas as brincadeiras, eu sei o quanto seu papel na família e na minha vida foram fundamentais. Você era uma argamassa meio Super Bonder que unia não só nosso núcleo familiar, como todas as 500 pessoas dessa grande família que é a nossa. Você faz muita falta e quero conseguir um dia distribuir bondade do jeito que você fazia. Nesse mestrado, muitas vezes onde eu achava que não tinha tempo ou energia, você insistiu para eu ir pra sua casa e isso recarregou minha motivação.

Às primas adolescentes, Clara e Zahra, obrigado por me ouvirem falando coisas nerds de psicologia e pelo menos fingirem que não acharam chato. Era bem comum eu perceber o quanto eu não sabia do que eu tava falando quando tentava explicar para vocês. Vocês são a energia da família! Aos primos Arthur, Léo e Helena, obrigado pelos momentos divertidos em família!

Vô, eu sei que não falo muito isso, mas você é um exemplo que eu uso para a minha vida. Sempre que tenho alguma dúvida ética sobre o que eu devo fazer, penso no que você faria no meu lugar. Agradeço por ser esse farol moral e por sempre se preocupar com o andamento do mestrado e da minha carreira.

Aos meus amigos Tropeço, Hugo, Pietro e David, valeu pelo apoio nesses 2 anos! Sei que eu meio que desapareci e quase não consegui encontrar vocês, mas vocês nunca desistiram de mim. Acredito que a amizade tenha pouco a ver com a distância e muito a ver com o vínculo. Obrigado pela paciência de sempre me ouvirem falando do meu projeto mirabolante e do apoio.

À equipe da Inframerica, obrigado também por não me deixarem desistir e por me ouvirem reclamando do quanto eu não estava dormindo. Agradeço muito pela companhia! Sarah, obrigado por me acompanhar nas quintas do energético, que em algumas épocas viraram quartas, terças, segundas... Obrigado por ter sido minha memória e consciência nos dias que eu estava com sono demais até para pensar.

Jacy, obrigado por ter acreditado em mim e ter me incentivado a seguir esse sonho que eu tinha de fazer mestrado. Você foi um dos motivos de eu ter acreditado que era possível e no final deu certo.

Professor Jairo, tenho que confessar que na primeira aula que tive com você, fiquei bem assustado. Eu tinha passado algumas horas lendo um texto que achava que tinha entendido, mas quando você começou a explicar ele, percebi que não tinha entendido nada. Seu conhecimento técnico é fenomenal e considero-me muito privilegiado de ter podido passar tanto tempo aprendendo com você. Além disso, sua humildade e forma de tratar as pessoas são exemplos que eu vou levar para a vida. Obrigado por ter acreditado em mim para executarmos um projeto tão fora do comum, mesmo eu não tendo muito tempo para dedicar ao mestrado.

Pai, foi um orgulho imenso para mim poder fazer pesquisa com você. Acho que poucos filhos têm um pai incrível assim e menos ainda podem fazer um mestrado trabalhando com ele. Apesar de sermos de áreas diferentes, nossa forma de pensar é muito parecida e foi bom demais poder discutir temas tão complexos e empolgantes com você. Obrigado por tudo que você me ensinou e me desculpe por tudo que eu não consegui aprender ainda. Obrigado também por todo o tempo que você gastou com a preparação e a análise dos dados mesmo com tanta coisa para fazer.

Mãe, acho que nenhum filho tem uma mãe que ajuda na análise de dados 1:30 da manhã quando acorda para beber água, mas eu tenho. Você viu toda a minha frustração e não só me incentivou o tempo todo como também me ajudou a estruturar métodos de pesquisa e análise

de dados. Sua dedicação e seu cuidado com a qualidade do seu trabalho me inspiram desde pequeno, quando eu te via trabalhando sábados e domingos de manhã. Se um dia eu conseguir ter metade da sua dedicação, já serei um profissional fora de série.

Professor Jairo, pai e mãe, vocês foram os gigantes em cujos ombros tentei subir para fazer esse trabalho.

SUMÁRIO

Lista de Tabelas.....	11
Lista de Figuras	14
Lista de Siglas	15
RESUMO	16
ABSTRACT	17
INTRODUÇÃO	18
CAPÍTULO 1 – MEDIDAS DO COMPORTAMENTO ORGANIZACIONAL: QUE FIZEMOS E PODEREMOS FAZER?.....	21
1.1 Introdução.....	21
1.2 Método.....	26
1.3 Resultados.....	28
1.4 Discussão	35
CAPÍTULO 2 – ESTILOS DE APRENDIZAGEM AUTODECLARADOS	38
2.1 Introdução	38
2.1.1 <i>Ambientes Virtuais de Aprendizagem</i>	38
2.1.2 <i>Aprendizagem no trabalho</i>	40
2.1.3 <i>Estilos de aprendizagem</i>	45
2.2 Método.....	52
2.2.1 <i>Construção da escala</i>	52
2.2.2 <i>Análises fatoriais exploratória e confirmatória</i>	56
2.3 Resultados.....	57
2.3.1 <i>Análise fatorial exploratória</i>	57
2.3.2 <i>Análise fatorial confirmatória</i>	64
2.4 Discussão	66
CAPÍTULO 3 – ESTILOS DE APRENDIZAGEM AUTODECLARADOS E OBSERVADOS	72
3.1 <i>Educational Data Mining e Learning Analytics</i>	73
3.2 Detecção Automática de Estilos de Aprendizagem.....	77
3.2.1 <i>Revisão de Literatura</i>	79
3.3 Criação de Ferramenta de Detecção Automática de Estilos de Aprendizagem	90
3.3.1 <i>Levantamento de Indicadores</i>	92
3.3.2 <i>Amostra de Cursos</i>	93

3.3.3 <i>Análise de Indicadores</i>	94
3.3.4 <i>Categorização dos Estilos de Aprendizagem Autodeclarados</i>	100
3.3.5 <i>Relação entre Indicadores e Estilos Autodeclarados</i>	103
3.3.6 <i>Teste da Detecção Automática de Estilos de Aprendizagem</i>	105
3.4 Comparação de Medidas	105
3.5 Discussão	108
DISCUSSÃO.....	113
REFERÊNCIAS	116
APÊNDICE A - Itens classificados por juízes	133
APÊNDICE B – Revisão de artigos de detecção automática de estilos de aprendizagem de 2015 a 2018	137
APÊNDICE C – Estrutura relacional do banco de dados	140
APÊNDICE D – Exemplo de relatório apresentado aos respondentes	144
APÊNDICE E – Análises descritivas da análise fatorial exploratória	145
APÊNDICE F – Lista de informações retiradas do banco de dados	148
APÊNDICE G – Tipos de recursos presentes no banco de dados	150
APÊNDICE H – Informações extraídas do banco de dados	151
APÊNDICE I – Indicadores criados a partir das informações extraídas do banco de dados .	154
APÊNDICE J – Comunalidades da análise fatorial exploratória dos indicadores de interação	156
APÊNDICE K – <i>Scree plot</i> da análise fatorial exploratória dos indicadores de interação	157

Lista de Tabelas

Tabela 1. Uso do método de observação relatado em publicações entre 1996 e 2017 por eixo paradigmático.....	34
Tabela 2. Instrumentos de estilos de aprendizagem utilizados em contexto de trabalho...	48
Tabela 3. Testes KMO e Bartlett da análise fatorial exploratória.....	58
Tabela 4. Comunalidades da análise fatorial exploratória sem retirada de itens.....	58
Tabela 5. Variância explicada e autovalores para a análise de Principais Componentes da análise fatorial exploratória.....	60
Tabela 6. Matriz de padrão da análise fatorial exploratória.....	61
Tabela 7. Testes KMO e Bartlett da análise fatorial exploratória após retirada de itens...	62
Tabela 8. Matriz de padrão da análise fatorial exploratória após a retirada de itens.....	63
Tabela 9. Correlação entre os fatores da análise fatorial exploratória após a retirada dos itens.....	63
Tabela 10. Variância explicada e autovalores da análise fatorial exploratória após a retirada de itens.....	64
Tabela 11. Resultados da revisão de literatura em Jaiswal et al. (2017).....	87
Tabela 12. Variância total explicada da análise fatorial exploratória dos indicadores de interação.....	98
Tabela 13. Matriz de padrão da análise fatorial exploratória dos indicadores de interação.	99
Tabela 14. Nomes e consistência interna dos fatores de indicadores de interação.....	99
Tabela 15. Quantidade de casos por <i>cluster</i>	101
Tabela 16. <i>Clusters</i> de cada um dos fatores de estilo de aprendizagem autodeclarado.....	103
Tabela 17. Correlações de Pearson entre os indicadores normalizados, os fatores de interação e as variáveis dependentes.....	104
Tabela 18. Teste de Omnibus da regressão logística dos dois fatores de estilo de aprendizagem autodeclarado e a quantidade de cursos realizados em relação à evasão.....	106
Tabela 19. Classificação de casos da regressão logística dos dois fatores de estilo de aprendizagem autodeclarado e a quantidade de cursos realizados em relação à evasão.....	106

Tabela 20. Classificação de casos usando <i>Random Committee</i> para os dois fatores de estilo de aprendizagem autodeclarado e a quantidade de cursos realizados em relação à evasão.....	107
Tabela 21. Teste de Omnibus da regressão logística dos quatro fatores de estilo de aprendizagem observado e a quantidade de cursos realizados em relação à evasão.....	106
Tabela 22. Classificação de casos da regressão logística dos quatro fatores de estilo de aprendizagem observado e a quantidade de cursos realizados em relação à evasão.....	107
Tabela 23. Classificação de casos usando <i>Locally Weighed Learning</i> para os dois fatores de estilo de aprendizagem autodeclarado e a quantidade de cursos realizados em relação à evasão.....	108

Lista de Quadros

Quadro 1. Categorias da validação por juízes.....	55
Quadro 2. Itens mantidos e retirados do fator teórico de Interação Passiva.....	67
Quadro 3. Itens mantidos e retirados do fator teórico de Processamento Cognitivo.....	68
Quadro 4. Itens mantidos e retirados do fator teórico de Regulação.....	69
Quadro 5. Itens finais do Fator 1.....	70
Quadro 6. Itens finais do Fator 2.....	71
Quadro 7. Matriz SWOT das pesquisas em <i>Learning Analytics</i> e <i>Educational Data Mining</i>	77
Quadro 8. Prós e Contras das abordagens colaborativa, baseada em dados e baseada na literatura.....	82
Quadro 9. Sistemas adaptativos de hipermídia em relação aos estilos de aprendizagem (Graf, 2007).....	85

Lista de Figuras

Figura 1. Quantidade de artigos publicados sobre CO e sobre escalas em CO por triênio.....	27
Figura 2. Quantidade de escalas publicadas em periódicos e em livros, por triênio.....	28
Figura 3. Percentual de escalas publicadas de 1996 a 2017 em livros e em periódicos, por temática.....	29
Figura 4. Usos de escala e de observação relatados em publicações entre 1996 e 2017, por temática.....	31
Figura 5. Modelo de Regulação dos Processos de Aprendizagem.....	48
Figura 6. <i>Scree plot</i> da análise fatorial exploratória.....	58
Figura 7. <i>Scree plot</i> da análise fatorial confirmatória.....	64
Figura 8. Modelo final da EEA-Ead_C.....	65
Figura 9. Áreas relacionadas a <i>Educational Data Mining</i>	72
Figura 10. Linha do tempo de acontecimentos relevantes de EDM.....	74
Figura 11. Relação entre abordagens e modelos de estilos de aprendizagem.....	80
Figura 12. Proporção média de indicadores por categoria de desvio-padrão.....	94
Figura 13. Distribuição de estudantes em relação ao percentual de indicadores que variaram mais de um desvio-padrão.....	95
Figura 14. Percentual de variações superiores a um desvio-padrão por indicador.....	96
Figura 15. Percentual de variações superiores a um desvio-padrão por fator.....	99
Figura 16. Gráfico de dispersão por fator de estilo de aprendizagem autodeclarado com classificação dos casos.....	101

Lista de Siglas

AVA – Ambiente Virtual de Aprendizagem

BI – *Business Intelligence*

CO – Comportamento Organizacional

DP – Desvio-padrão

DBD – Detecção Baseada em Dados

DBT – Detecção Baseada em Teoria

EDM – *Educational Data Mining*

EAD – Ensino a Distância

EEA-Ead_C – Escala de Estilos de Aprendizagem em Ensino a Distância com base em Comportamentos

ILS – *Index of Learning Style*

ISD – Instructional System Design

LA – *Learning Analytics*

LAK – *Learning Analytics and Knowledge*

ME – Modelo do Estudante

POT – Psicologia Organizacional e do Trabalho

RH – Recursos Humanos

SEAH – Sistema Educacional Adaptativo em Hipermedia

RESUMO

A análise de custo-benefício, entre generalidade, precisão e realismo, guia escolhas metodológicas em comportamento organizacional (CO). O *big data* e o *machine learning* podem tornar obsoleta essa análise e ajudar a subárea a focar novamente o “fazer” e não apenas o “dizer”. O uso de observação de comportamento foi sistematicamente revisto em artigos e livros (1996-2017) brasileiros. Dentre 1.062 artigos com relatos de pesquisa sobre CO, publicados em 19 periódicos de psicologia e administração, 68% usaram escalas e apenas 10% a observação. A presente pesquisa objetivou propor um método para comparar medidas autodeclaradas (escalas) com comportamentos observados usando *machine learning*. Foi escolhido o construto de estilo de aprendizagem para propor o método, que analisou de forma comparativa e complementar a relação entre o que é declarado e o que é observado. Os estilos de aprendizagem autodeclarados foram medidos por meio de uma nova escala, que foi elaborada e aplicada em 1.764 estudantes de um ambiente virtual de aprendizagem (AVA) de um órgão da administração federal. Para a escala, foram encontradas evidências de validade exploratória e confirmatória, sendo a estrutura final composta por 18 itens agrupados em dois fatores. As informações brutas do AVA (*footprints*) foram transformadas em indicadores, que foram comparados com os estilos identificados por meio de questionário e também agrupados em fatores.. Foi encontrado que os padrões comportamentais dos estudantes, referentes aos estilos, são relativamente constantes em diferentes cursos. Além disso, foram encontradas evidências de validade para uma estrutura de três fatores de indicadores referentes a interação. Os estilos observados e os autodeclarados foram testados em relação à evasão e ambos apresentaram um poder de predição relevante. A triangulação metodológica de observação de comportamento e escalas foi mais efetiva quando usados de forma complementar do que de forma comparativa. O estudo apresentou uma alternativa metodológica inovadora e viável para a análise de estilos de aprendizagem em AVA, usando *machine learning*.

Palavras-chave: análise multivariada, observação de comportamento, *big data*, estilos de aprendizagem, *machine learning*.

ABSTRACT

Tradeoff analysis, between generalizability, precision and realism, guides methodological choices in organizational behavior (OB). Big data and machine learning can make this tradeoff analysis obsolete and help OB to once again focus on “doing” and not only on “saying”. Use of the behavior observation method was systematically reviewed in Brazilian articles and books (1996-2017). Among 1,062 research articles about OB, published in 19 psychology and business journals, 68% used scales and only 10% used observation. The goal of this research is to propose a method to compare self-declared measures (scales) with observed behaviors using machine learning. The method analyzes what is observed and what is declared in opposite and complementary approaches. Learning styles were used to develop and apply the method. A new scale, applied in 1,764 students in a virtual learning environment (VLE), was developed to identify self-declared learning styles. Evidence for the exploratory and confirmatory validity and reliability is presented for a structure with 18 items grouped in two factors. Raw data from the VLE (footprints) were transformed indicators. Those indicators were compared to the responses to the scale and also grouped in factors. The analysis of the indicator showed behavior patterns related to learning styles were relatively constant between courses and four factors of interaction indicators were identified. Observed and self-declared learning styles were separately tested in their prediction of dropout. Both were able to significantly predict dropout. Methodology triangulation of behavior observation and scales was more effective when used in a complementary approach. This study presents a viable and innovative methodological alternative for analyzing learning styles in VLE, using machine learning

Keywords: multivariate analysis, behavior observation, big data, learning styles, machine learning.

INTRODUÇÃO

As pesquisas em Comportamento Organizacional (CO) analisam quase que exclusivamente o que os indivíduos falam e não o que eles fazem. Como exemplo disso, apenas 12,4% dos estudos brasileiros da subárea usam o método de observação do comportamento (Mourão, Bastos & Oliveira, 2016). A preferência metodológica pelas escalas (Schein, 2015) é justificada pela dificuldade de observar o comportamento dos indivíduos. As principais dificuldades para a aplicação do método de observação são: (i) abertura da organização; (ii) confiabilidade dos dados (iii) consistência dos dados; e (iv) tempo gasto pelos observadores para analisar apenas uma pequena fração dos comportamentos (Mourão *et al.*, 2016).

Inovações tecnológicas como o *big data* e o *machine learning* estão entre as tendências para a Psicologia Organizacional e do Trabalho (POT) (Grimm, Jacobucci & McArdle, 2017). As organizações guardam imensas quantidades de dados sobre comportamentos em forma bruta. Apesar de serem inúteis nessa forma, técnicas computacionais de análise de dados como o *data mining* permitem que os dados sejam transformados em informações relevantes (Guzzo, Fink, King, Tonidandel & Landis, 2015). Um exemplo de plataforma que armazena dados e que está presente em organizações são os Ambientes Virtuais de Aprendizagem (AVAs).

O construto estilo de aprendizagem é complexo e tem quase tantas definições quanto autores (De Bello, 1990). Ele pode ser entendido como a combinação de características cognitivas, afetivas e fatores psicológicos que se expressam de forma relativamente estável na maneira como o aprendiz percebe, interage com e responde a ambientes de ensino (Keefe, 1979 em Feldman, Monteserin & Amandi, 2015). AVAs são ambientes de ensino. Comportamentos podem ser usados para detectar os estilos de aprendizagem comparando-os

com a resposta a escalas de estilos (Al-Azawei & Badii, 2014; Feldman *et al.*, 2015; Oliveira, Sales, Pereira & Moreira, 2018; Jaiswal, Singh & Ahuja, 2017). No entanto, a detecção de estilos dessa forma é baseada apenas nas escalas e sua construção é específica para um ou poucos cursos. Uma das formas de suprir essa lacuna é comparar os estilos baseados em respostas a escalas (autodeclarados) com padrões comportamentais (observados) em AVAs.

Tendo em vista esse cenário, o objetivo geral desta Dissertação é propor um método de comparação entre estilos de aprendizagem autodeclarados e observados em diferentes cursos de um AVA. Para que ele seja atingido, foram definidos quatro objetivos específicos:

1: Elaborar uma escala de estilo de aprendizagem autodeclarado, com bons indícios de validade estatística, adequada para ser observada em AVA.

2: Identificar padrões comportamentais relativamente estáveis e observáveis em diferentes contextos de AVA.

3: Detectar os estilos de aprendizagem autodeclarados por meio de comportamentos observáveis em AVA.

4: Comparar os estilos de aprendizagem autodeclarados e observados, no que concerne à predição de evasão de cursos em AVA.

O Capítulo 1 analisou de forma crítica as escolhas metodológicas de (CO), focando no uso de medidas. Para isso, inicialmente foi feita uma revisão de literatura que abarcou os usos de escalas e da observação do comportamento na pesquisa brasileira. Em seguida, foram sugeridas soluções para o uso de tecnologia para a observação direta do comportamento, com base na produção científica internacional.

O Capítulo 2 descreve a criação da Escala de Estilos de Aprendizagem em Ensino a Distância com base em Comportamentos (EEA-Ead_C), visando alcançar o objetivo específico 1. Para isso, foram analisadas as literaturas de aprendizagem no trabalho e estilos de aprendizagem, a fim de justificar as escolhas teóricas e metodológicas de criação da escala.

Em seguida, foi exposto o método usado para a criação da escala, assim como para as análises exploratória e confirmatória realizadas. Por último, foram apresentados os resultados do processo de validação e eles são discutidos.

No Capítulo 3, foram testadas as relações entre estilos de aprendizagem autodeclarados e observados. Os estilos autodeclarados foram obtidos a partir das respostas às escalas e os estilos observados por meio de comportamentos observados em treinamentos oferecidos em AVA. Para fazer isso, primeiro foram apresentadas as áreas de *Educational Data Mining* e *Learning Analytics* e em seguida foi realizada uma revisão de literatura sobre detecção automática de estilos de aprendizagem usando comportamentos em AVA. Quanto aos estilos autodeclarados, os participantes dos treinamentos (estudantes) foram agrupados com base nas suas respostas ao questionário. Quanto aos estilos observados, foram elaborados indicadores a partir do banco de dados, que foram analisados em relação a estudantes que participaram de mais de um treinamento (objetivo específico 2), e foram agrupados com base na teoria e em análises estatísticas. Buscou-se criar uma ferramenta que identificasse os estilos autodeclarados a partir dos estilos observados e dos comportamentos (objetivo específico 3). Por fim, os estilos de aprendizagem autodeclarados e os observados foram comparados em relação a seu poder de predição da evasão nesses treinamentos oferecidos em AVAs (objetivo específico 4) e os resultados foram discutidos.

CAPÍTULO 1 – MEDIDAS DO COMPORTAMENTO ORGANIZACIONAL: QUE FIZEMOS E PODEREMOS FAZER?

1.1 Introdução

Mudanças geopolíticas e econômicas e inovações tecnológicas promoveram alterações significativas nas organizações e no trabalho. Essas alterações impactaram a subárea da POT, que cresceu e se consolidou. Como subárea do conhecimento, o objeto de estudo da POT é o fazer humano, a ser investigado em contextos do trabalho e das organizações. Ela também pode ser compreendida como campo de aplicação, no qual precisa resolver problemas práticos relativos à interação entre o comportamento humano, o trabalho e as organizações e relativos às práticas dessas organizações. Mas, estará aquela subárea do conhecimento, quando decide sobre como investigar, aproveitando* todo o potencial das inovações tecnológicas presentes neste campo de aplicação? O presente capítulo analisou os métodos de investigação usados em POT, por meio de uma revisão sistemática da produção científica brasileira.

Estudar a influência mútua entre indivíduos e organizações é um desafio, que necessita de explicações advindas dos níveis individual, das equipes e organizacional. Antropologia, sociologia, administração e psicologia buscaram estudá-la a partir de seus arcabouços teóricos e metodológicos, ou criaram novos arcabouços. Aquela influência foi recortada em construtos e relações analisadas por diversos ângulos e abordagens metodológicas. A forma ampla e difusa dessa investigação dos fenômenos levou a uma grande fragmentação (Schein, 2015). A POT assim construída tem muitas subculturas, cada uma com seus próprios jargões e procedimentos. Cada uma desenvolvendo um silo intelectual distante de uma teoria central e de outros silos, segundo este autor. Outro efeito foi a desconexão entre o que a academia produz (área de conhecimento) e o que os profissionais

* O Capítulo 1 dessa Dissertação foi submetido na forma de artigo a um periódico com aprovação ainda pendente.

utilizam (campo de aplicação). Pesquisas são desenhadas para obterem o máximo possível de validade e não necessariamente para obterem máximo impacto. Por outro lado, profissionais podem distorcer os achados dos acadêmicos para justificarem suas intervenções.

O cenário brasileiro do campo de aplicação da POT iniciou com forte presença do uso de medidas sem evidências de validade no país (Borges-Andrade, Pérez & Toro, 2018). A quantidade de profissionais neste campo e os programas de pós-graduação, que realizam a produção que constitui a subárea de conhecimento, aumentaram substancialmente. O crescimento acelerado da subárea de conhecimento culminou em uma grande quantidade de publicações de abordagens diversas, segundo estes autores. A formação em nível de graduação conta com poucas disciplinas de POT e os estágios profissionais, além de muitas vezes não serem obrigatórios no campo de aplicação da POT, focam na apreensão e na aplicação. Não são desenvolvidas competências ligadas à criação de conhecimento e avaliação de intervenções. Este cenário resultou em uma formação acadêmica inadequada, tornando a prática profissional (campo de aplicação) pouco crítica e desconectada da produção acadêmica (subárea de conhecimento), segundo Borges-Andrade *et al.*, (2018).

A POT pode ser organizada em três eixos no Brasil, com base nos paradigmas que adotam: comportamento, subjetividade e clínico (Bendassoli, Borges-Andrade & Malvezzi, 2010). O primeiro tem como base o pós-positivismo e busca a objetividade, mesmo entendendo que não é possível alcançá-la completamente. É influenciado pelo *behaviorismo*, pelo *neobehaviorismo* e pelo sociocognitivismo, que originalmente utilizaram métodos experimentais. O segundo parte do paradigma crítico, que contrapõe o pós-positivismo, em seu planejamento, execução e interpretação dos resultados de pesquisa. É influenciado pelas abordagens sócio histórica e institucional e pelo pós-estruturalismo. Usa métodos e técnicas que focam em uma análise do fenômeno e deixam em segundo plano a comparabilidade e a generalização dos resultados. O terceiro também tem como paradigma teórico o crítico e

adota metodologias qualitativas para a análise histórica do indivíduo. Inclui a ergologia, a psicodinâmica do trabalho, a sociologia clínica e a clínica da atividade.

O principal ponto de tensão, principalmente no eixo do comportamento, é a relação entre a relevância e o rigor (Bendassoli *et al.*, 2010). Quanto maior o rigor metodológico (e.g., desenhos experimentais isolando variáveis), maior o distanciamento do contexto onde o fenômeno se manifesta e menor a aplicabilidade dos resultados. Somando isto ao alto custo dos desenhos experimentais, aumentou intensamente o uso de *surveys* (correlacionais). Ocorreu ainda aumento de pesquisas qualitativas. Nos últimos anos, muitos periódicos passaram a pressionar pelo uso de delineamentos de pesquisa mais robustos. Por outro lado, os profissionais da POT frequentemente recorrem a soluções de fácil aplicação, mas de validade e consistência duvidosas, segundo estes autores.

Apesar do agrupamento das diferentes vertentes de POT em 3 eixos, no interior desses eixos não há consistências teórica e metodológica. Neles, pesquisas partem de diferentes perspectivas e analisam os mesmos fenômenos e relações complexas. Isto culminou em grande proliferação de construtos, que sofrem principalmente de quatro problemas de definição: (i) alongamento ou combinação; (ii) sobreposição; (iii) tautologia; e (iv) apropriação e domesticação (Rodrigues & Carvalho-Freitas, 2016). A sobreposição de conceitos e a fragmentação teórica impedem que sejam descobertas a natureza e a intensidade das relações com outros conceitos e variáveis. Uma vez que as diferentes formas de investigar os fenômenos não são integradas, os conhecimentos gerados em cada um dos eixos não são cumulativos.

CO é um campo apoiado pela psicologia, sociologia e economia. Tem as atividades organizacionais como objeto, sintetizadas em três níveis distintos e interligados: macro (organizações), meso (equipes e grupos) e micro (indivíduo), segundo Siqueira (2009). CO adotou a ideia de que a forma de evoluir uma teoria seria diferenciar o comportamento em

múltiplos componentes, medir esses componentes e, por análise fatorial, descobrir as variáveis subjacentes (Schein, 2015). Uma produção que se baseia em uma categoria de método não está se esquivando dos erros de medida. Ela acaba privilegiando um tipo de erro. A visão dilemática da pesquisa em psicologia sugere que não há apenas um método ou um conjunto de métodos que garantem o sucesso. Esta visão aponta que todos os métodos têm erros e que os pesquisadores precisam aprender a conviver com isso. Seguindo esta óptica, há três extremos (ou chifres): a generalidade, a precisão e o realismo. Quando se escolhe privilegiar um dos chifres, os outros serão prejudicados. Quando se escolhe dois deles, o terceiro também o será. Porém, não seria possível privilegiar os 3 chifres simultaneamente. Este dilema sugere, como solução, a diversificação metodológica e um grande volume de pesquisas para se estudar um mesmo fenômeno (McGrath, 1981).

A pesquisa científica internacional de POT quase sempre se baseia apenas no que o indivíduo fala, por meio de análises qualitativas ou quantitativas. Esta forma predominante de acesso ao “dizer” coloca em questão a validade dos achados (Aguinis, Pierce, Bosco & Muslin, 2009). Apenas 12,4% dos estudos brasileiros relataram observação como método e a maior parte deles também é focada no que o indivíduo diz, apesar da pretensão de considerar o que as pessoas fazem, não incorporando elementos fundamentais para a apreensão apropriada do fenômeno (Mourão, Bastos & Oliveira, 2016). Os estudos de desempenho, comprometimento e cidadania organizacional, que apesar de serem comportamentos, são investigados quase que exclusivamente pelo que os indivíduos dizem, são exemplos mencionados por estes autores.

O modelo de resposta a escalas, especialmente as referentes a comportamentos, aponta passos que supostamente o respondente siga ao responder um item: (i) entender a pergunta; (ii) compreender o comportamento de interesse; (iii) identificar corretamente o período de interesse; (iv) lembrar desse período para identificar todas as ocorrências relevantes do

comportamento; (v) datar corretamente essas ocorrências lembradas; (vi) somá-las corretamente; (vii) encaixar a quantidade delas nas alternativas de respostas propostas; e (viii) responder a pergunta (Schwarz & Oyserman, 2001). Cada um desses passos abre espaço para erros e, mesmo a medida tendo consistência interna, não é possível atestar sua validade externa. Para apreender os construtos da POT, considerando a complexidade inerente a eles e a dificuldade de categorizá-los de forma válida e confiável, é necessária uma triangulação de métodos. Nela, devem ser usadas diferentes formas de investigar o “fazer” e o “dizer”, de maneira que as informações sejam usadas em conjunto. Os estudos devem tirar um pouco o foco nas respostas baseadas apenas no “dizer”, e buscar também formas de acesso ao “fazer”, sugerem Mourão *et al.* (2016). Isso provavelmente envolveria a observação direta.

Existem explicações históricas, limitações tecnológicas e dificuldades operacionais para usar o método de observação direta do trabalho em organizações. A primeira é conseguir que estas permitam a observação de seus funcionários. A segunda é a confiabilidade dos dados coletados, pois a presença de observador afeta o desempenho de quem é observado. O terceiro é o custo-benefício, pois o método demanda muito tempo e mais de um observador, para acesso a uma pequena fração do comportamento no trabalho. O quarto é que a consistência dos dados é dificilmente atingida, pois observadores precisam estar bem treinados. Por último, há variáveis de contexto que não podem ser controladas, culminando em baixa reprodutibilidade (Mourão *et al.*, 2016). Os métodos de observação foram quase inteiramente substituídos no eixo de comportamento por *surveys* e são usados de forma assistemática nos outros dois eixos da POT, como suporte à investigação do contexto do indivíduo. É preocupante que uma subárea do conhecimento que objetiva explorar, prever e correlacionar comportamentos tenha supostamente abandonado o acesso direto a eles. Mas há uma mudança em curso.

O grande e variado volume de dados atualmente disponíveis está revolucionando a economia e as ciências sociais aplicadas. Métodos de *big data* podem identificar sinais complexos nesses dados, com base em técnicas estatísticas (*machine learning*, *statistical learning* e *data mining*), afirmam Grimm, Jacobucci e McArdle (2017). Seu caráter exploratório e a crença de que não temos volumes de dados suficientemente grandes limitou seu uso na pesquisa em psicologia, embora *data mining* devesse ser usado para aprendermos sobre esses dados e apoiar a geração de hipóteses, argumentam estes autores. Tais métodos tem o potencial de eliminar, ou pelo menos reduzir, dilemas como o dos três chifres, antes tidos como inerentes à investigação de fenômenos em ciências sociais. Generalidade, precisão e realismo poderiam ser simultaneamente alcançadas. A investigação científica passaria a ser preditiva ao invés de explicativa (Chang, Kauffman & Kwon, 2013). No contexto de POT, estes dados podem estar disponíveis em arquivos organizacionais imensos com informações sobre comportamentos dos funcionários. Os desafios que trazem, além dos éticos, são a integridade e a análise de dados. As informações gravadas podem ser irrelevantes e são necessárias técnicas computacionais específicas para agrupar e analisar os dados. As principais tendências de uso de *big data* em POT são em seleção de pessoas, avaliação de desempenho e identificação de redes de relacionamentos (Guzzo, Fink, King, Tonidandel & Landis, 2015). Considerando essa nova oportunidade e aqueles dilemas, o Capítulo 1 analisará os métodos utilizados nas publicações brasileiras sobre CO.

1.2 Método

A produção brasileira da área de CO foi analisada usando o método descrito por Bastos, Maia, Rodrigues, Macambira e Borges-Andrade (2014), Cantal, Borges-Andrade e Porto (2015), Fonseca, Porto e Borges-Andrade (2015), Nascimento, Borges-Andrade e Porto (2016), Resende e Neiva (2018) e Sampaio, Borges-Andrade e Bonatti (2018). Neste método, estudantes de pós-graduação, previamente treinados em análise de relatos de pesquisa sobre

CO, fazem o levantamento de artigos em duplas e outras duas duplas analisam, após confirmar o acerto dos artigos levantados. Ensaios, revisões de literatura e resenhas são descartados. Somente artigos que relatam estudos empíricos com variável critério de CO em nível micro, realizados com trabalhadores brasileiros, são considerados. Os artigos são classificados em 17 grandes temáticas de comportamento organizacional: (i) Afeto; (ii) Aprendizagem; (iii) Atitudes frente a mudança; (iv) Clima organizacional; (v) Cognição; (vi) Competências; (vii) Comportamentos de fuga e esquiva; (viii) Contratos psicológicos; (ix) Criatividade e solução de problemas; (x) Cultura organizacional; (xi) Desempenho produtivo; (xii) Desvios de comportamento; (xiii) Bem-estar e saúde; (xiv) Interações sociais; (xv) Motivação; (xvi) Significados, sentidos e identidade; e (xvii) Tomada de decisão. Algumas ainda são divididas em subtemáticas. Em cada artigo, são coletadas informações sobre a natureza e finalidade do estudo, características do desenho de investigação, amostragem, origem dos dados, tipo de medida, procedimento de análise dos dados, setor e segmento da economia estudados. Os resultados desse trabalho são registrados em planilha.

Para a presente revisão, foram escolhidos os 18 periódicos de administração (Revistas de Administração da USP e Mackenzie, Revista de Administração Contemporânea e Brazilian Administration Review da ANPAD, Revistas de Administração Pública e de Empresas da FGV, Revista Organizações e Sociedade e Revista Eletrônica de Administração) e psicologia (Estudos de Psicologia da UFRN e da PUC-Campinas, Psicologia Reflexão e Crítica, Psicologia em Estudo, Psicologia Teoria e Pesquisa, Revistas Psico da USF e da PUC-RS, Psicologia Ciência e Profissão, Avaliação Psicológica, e Revista Psicologia Organizações e Trabalho), mencionados por Sampaio *et. al* (prelo), e mais o periódico Paidéia, de psicologia. Neles, foram levantados 1062 artigos, publicados de 1996 a 2017, que atendiam os critérios anteriormente descritos. Os instrumentos de coleta de dados relatados foram classificados como: Escala ou questionário, Observação, Teste ou prova, Outro. Foram analisados quanto

ao eixo paradigmático do estudo, ao uso ou não de análise quantitativa, à definição prévia de comportamentos ou construtos a serem observados, à sistematicidade da observação, à participação do pesquisador e à forma como foram usados os resultados na pesquisa. As observações realizadas nos estudos foram categorizadas primeiramente quanto a sua estrutura, podendo ser exploratórias, com construto definido ou com comportamentos definidos. A segunda categorização foi dos resultados, que poderiam não ter sido citados, ter sido usados apenas para levantar informações para outras ferramentas ou ter sido usados para o próprio objetivo do artigo. A terceira categorização foi quanto à amostragem, que poderia ter sido um estudo de caso, um estudo com amostra de apenas uma organização ou um estudo com amostra de várias organizações. A quarta e última categoria foi de participação do pesquisador, que poderia ter sido participante ou não-participante. Provas e testes e outros instrumentos de coletas de dados foram desconsiderados, pois representaram menos de 5% das publicações em periódicos. Além desses 19 periódicos, foram analisados os quatro livros de escalas de POT publicados nos últimos 22 anos (Abbad *et.al.*, 2009; Siqueira, 2009; Siqueira, 2014; Puente-Palacios & Peixoto, 2015). Cada uma dessas escalas (n=87) em cada um desses livros foi classificada de acordo com as mesmas 17 temáticas em que foram classificados os artigos.

1.3 Resultados

Houve dois momentos de aceleração na divulgação da produção nacional sobre CO, um no triênio de 2000 a 2002 (107%) e outro no triênio de 2009 a 2011 (35%). Quanto às escalas, quase inexistentes até o início dos anos 2000, houve três momentos de aceleração, sendo um de 2000 a 2005 (333%), outro de 2009 a 2011 (56%) e o último de 2015 a 2017 (89%). A publicação de escalas acompanhou o padrão de crescimento da publicação em CO (Figura 1).

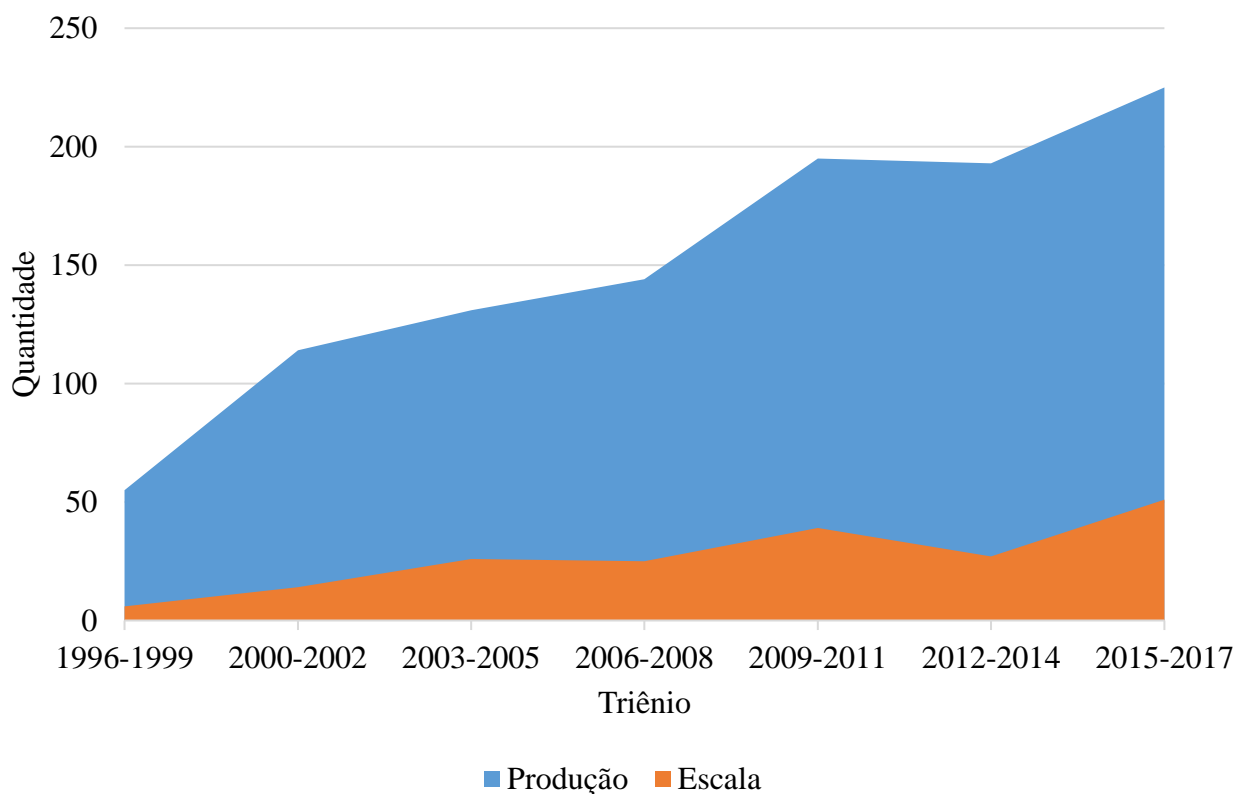


Figura 1. Quantidade de artigos publicados sobre CO e sobre escalas em CO por triênio

Um dos marcos para a produção de escalas foi o artigo de Siqueira (2002) que, ao analisar as medidas de CO até 2002, apontou a ausência de padronização no processo de construção de medidas e a falta de medidas com boas evidências de validade. Destacou que a diversidade teórica dificultava a definição de construtos a serem analisados e também de procedimentos de análise que realmente captassem toda a complexidade dos fenômenos. Em última instância, isto culminava no uso de escalas com baixa evidência de validade ou de escalas que foram construídas para outros contextos. Houve um grande crescimento da produção de escalas desde 2002. No entanto, esse crescimento não representa uma unificação metodológica ou procedimental, como proposto pela autora, e sim uma fragmentação ainda maior da subárea. À medida que os eixos paradigmáticos de CO passaram a aceitar cada vez mais as escalas como um instrumento de medida, cada um deles desenvolveu instrumentos para medir seus próprios construtos, ao invés de buscarem uma unificação. Na grande maioria

dos estudos (87%) onde há a construção de uma escala, o objetivo final é só a própria construção da escala. Comparando as escalas publicadas em livros e em artigos, a produção foi parecida de 1996 a 2007, mas tornou-se muito discrepante na última década. A produção de escalas relatada em artigos nos últimos 10 anos foi mais do que o dobro da produção de escalas em livros (Figura 2).

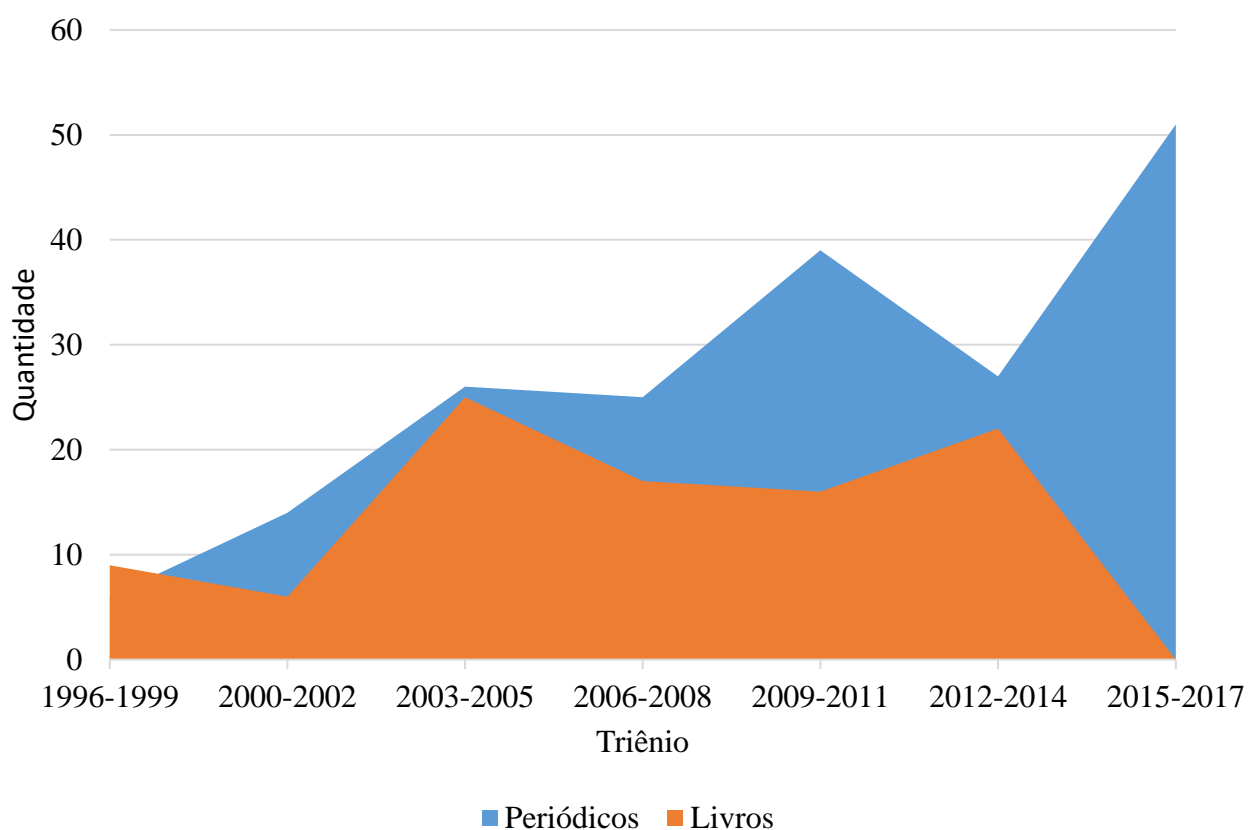


Figura 2. Quantidade de escalas publicadas em periódicos e em livros, por triênio.

As temáticas de CO onde há menos produção de escalas são: tomada de decisão e julgamentos; criatividade e solução de problemas; comportamentos de fuga e esquivas; e atitudes frente a mudanças (Figura 3). Dentre as temáticas em que a produção é relevante, as que são equilibradas quanto aos dois meios de divulgação são: significados, sentidos e identidade; motivação; interações sociais; desvios de comportamento; competências;

aprendizagem; e afeto. As temáticas que seguem o padrão de maior divulgação de escalas em periódicos do que em livros são: bem-estar e saúde; desempenho produtivo; e clima organizacional. Por último, as temáticas que são opostas ao padrão, com mais relatos em livros do que em periódicos são: cultura organizacional; contratos psicológicos e cognição.

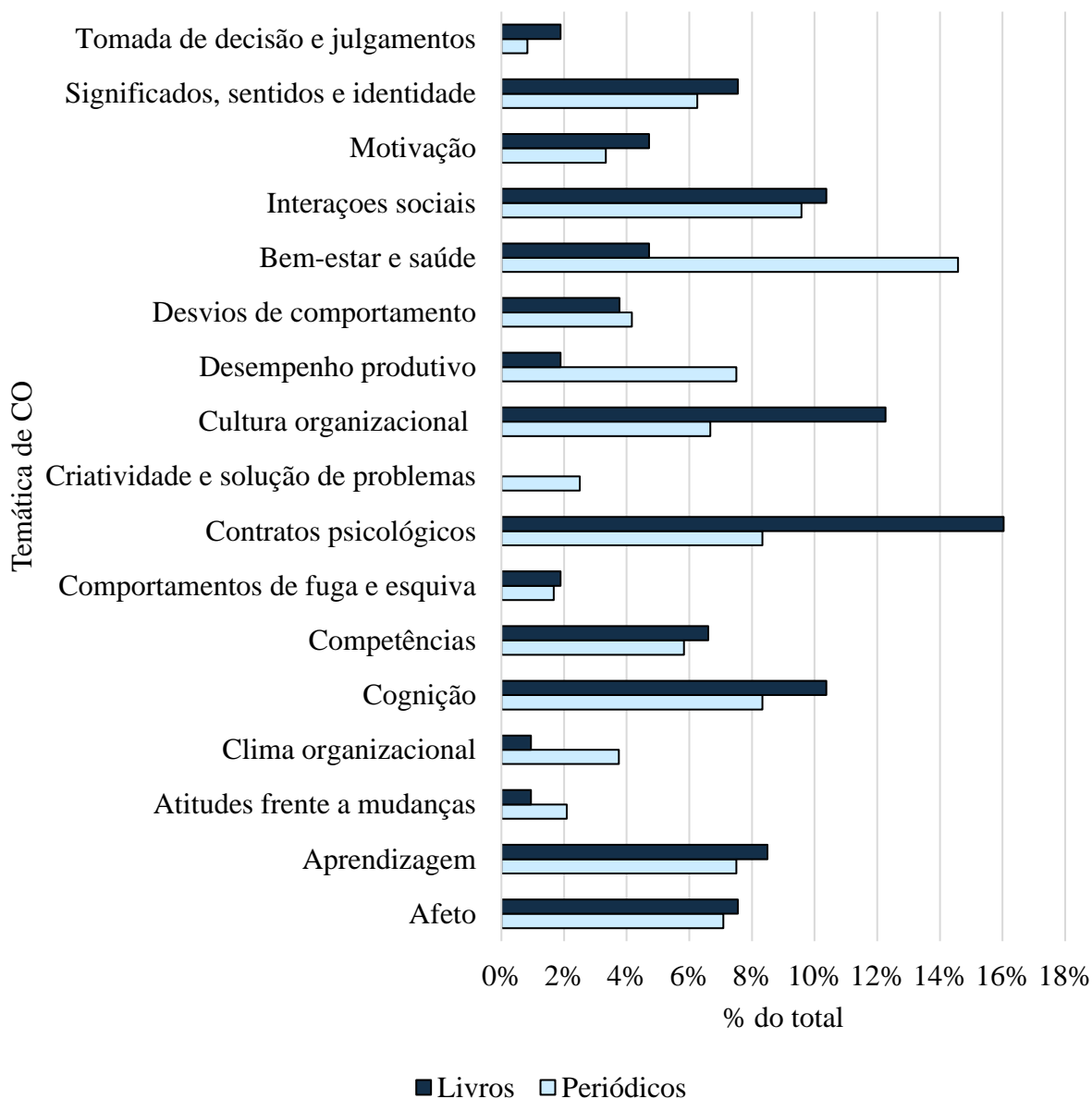


Figura 3. Percentual de escalas publicadas de 1996 a 2017 em livros e em periódicos, por temática.

Os dados são referentes ao volume de relatos sobre escalas e não ao seu uso. Por exemplo, escalas de clima organizacional estão entre as mais utilizadas por profissionais de

POT, embora se saiba que muitas delas nunca passaram por qualquer análise sobre evidências de validade, tendo sido produto de soluções “caseiras” ou diretamente importadas por empresas de consultoria que vendem serviços de diagnóstico de clima organizacional. A quantidade de escalas de clima em livros é muito pequena em relação ao total de escalas em livros (0,9%), sendo menos representativa inclusive que a quantidade de escalas em artigos (3,8%).

O uso da observação como método de coleta de dados, em estudos sobre CO, é escasso (8%). Por outro lado, as escalas estão presentes em boa parte desses estudos (68%). A representatividade do uso de escalas varia muito de acordo com a temática de CO, sendo que, por exemplo, quase todos os estudos de contratos psicológicos contam com o uso de uma escala e menos da metade dos estudos de significados, sentidos e identidade contam com o uso de escalas (Figura 4). Apesar desta variação, todas as temáticas têm um número significativo de estudos onde escalas são usadas. Ao contrário, o uso do método de observação varia pouco entre as temáticas de CO. Mesmo temáticas que se referem diretamente a comportamentos, como comportamentos de fuga e esquiva e desvios de comportamentos, têm poucos ou nenhum estudo que relataram o uso do método de observação.

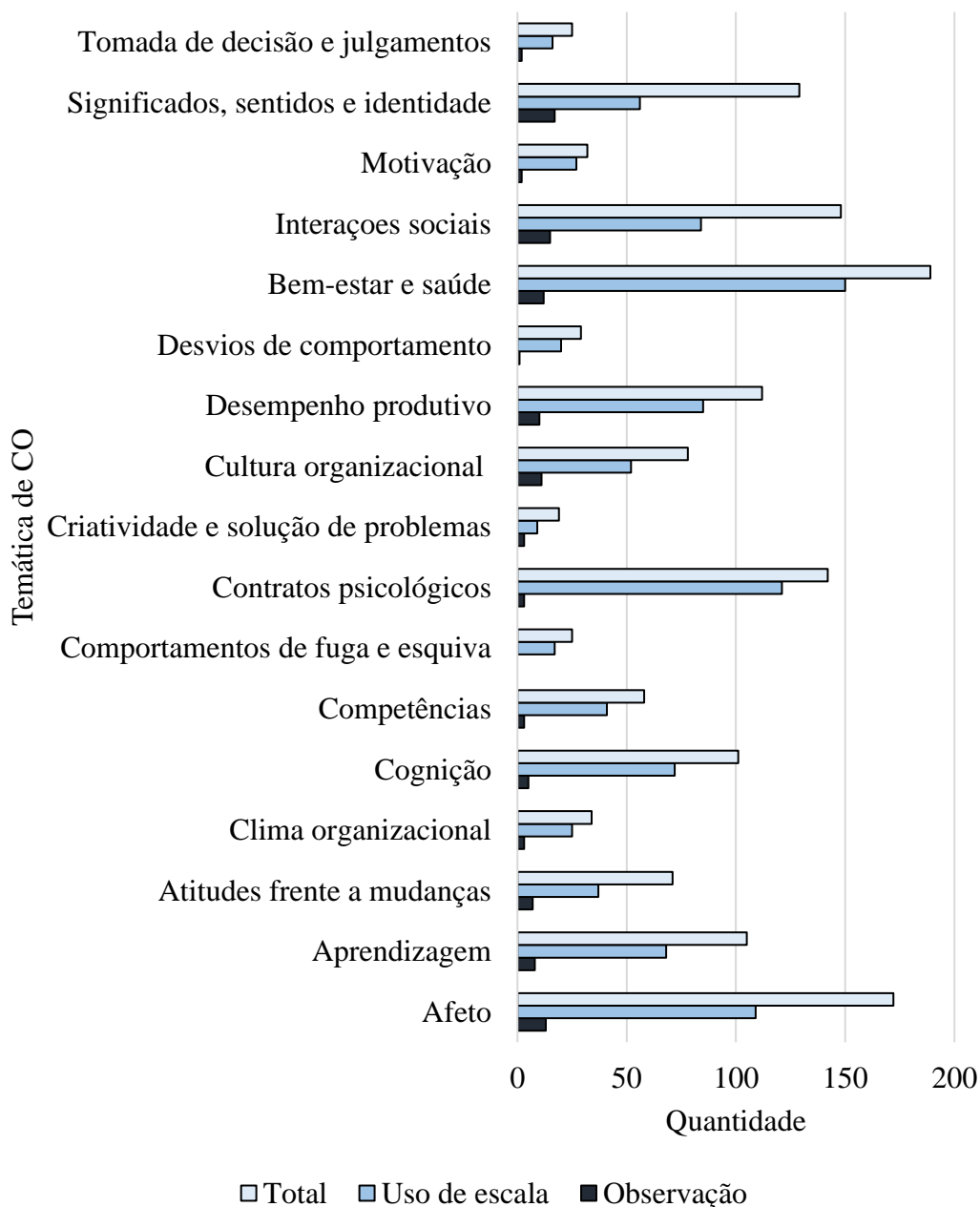


Figura 4. Usos de escala e de observação relatados em publicações entre 1996 e 2017, por temática.

Uma análise mais aprofundada dos 81 artigos encontrados, no qual o método de observação foi relatado, revela alguns possíveis padrões. O primeiro é que quase nenhum deles apresentou uma análise quantitativa dos dados de observação. Apesar de alguns estudos usarem combinações de métodos como o diário de campo e a análise categórica de Bardin, o que permitiria uma análise quantitativa de dados, nenhum descreveu tais análises. Apenas 3 dos 81 usaram algum tipo de análise quantitativa. Representam uma exceção à regra que

parece ter dominada o uso de observação nas pesquisas sobre CO encontradas na presente revisão sistemática. Além disso, nenhum deles utilizou métodos de *big data*, como *data mining*.

A maioria dos artigos que descrevem o uso do método de observação é do eixo da subjetividade e muito poucos são do eixo clínico (Tabela 1). Quanto aos procedimentos de observação adotados, a grande maioria dos estudos usa o método exploratório, sem definir os comportamentos a serem observados e nem o construto de interesse para o qual deveriam ser observadas expressões por meio de comportamentos. Há uma grande variedade de métodos que usam a observação de forma exploratória (e.g., etnografia), no entanto, estes métodos buscam entender uma realidade específica e seus resultados não podem ser generalizados. Quanto aos resultados da observação, no eixo comportamental, em mais de um quarto dos casos eles não são sequer citados. Quando citados, muitas vezes eles serviram apenas de suporte para a criação de questionários ou escalas. Isto é, os achados obtidos por meio de observação apenas serviram em uma etapa intermediária de um processo de pesquisa, em que um levantamento sobre o “fazer” seria substituído pela investigação do “falar”.

Tabela 1

Uso do método de observação relatado em publicações entre 1996 e 2017 por eixo paradigmático.

	Quantidade	Resultados		
		Não citados	Auxiliam outras ferramentas	Usados para o objetivo do artigo
Clínica	4	0%	0%	100%
Comportamental	28	25%	11%	64%
Subjetividade	49	6%	8%	84%
Total	81	12%	9%	77%

1.4 Discussão

A produção brasileira de CO, principalmente do eixo do comportamento, pode estar privilegiando exageradamente as escalas como método de coleta de dados. Muitos artigos são publicados apenas com o relato da validação de uma escala. Considerando que os profissionais da área de aplicação da POT têm como base os livros e não os artigos (Borges-Andrade *et al.*, 2018), podem existir muitas escalas validadas que são pouco ou nada usadas. Os construtos para os quais as escalas foram criadas não necessariamente podem ser úteis para os profissionais da área, já que há uma discrepância significativa entre as temáticas das escalas publicadas em periódicos e em livros.

As escalas privilegiam a generalidade. Mas podem carecer de precisão, uma vez que estão sujeitas aos problemas de fidedignidade apontados por Schwarz e Oyserman (2001) e aos problemas de validade, por se basearem unicamente no que o indivíduo diz (Mourão *et al.*, 2016). Muitos construtos, apesar de serem investigados em diversos contextos, não são necessariamente fiéis à realidade. É necessária uma triangulação de métodos para diminuir os problemas de medida. O eixo paradigmático que mais usa métodos qualitativos e que também tem maior diversidade metodológica em seu uso é o da subjetividade. No entanto, os outros dois eixos quase não se beneficiam dessas investigações, pois as bases teóricas e os conceitos usados pelos eixos são diferentes, apesar dos fenômenos observados serem similares (Rodrigues & Carvalho-Freitas, 2016). Considerando que a diferenciação dos eixos paradigmáticos se estende às escolhas metodológicas, todos eles têm lacunas em suas investigações.

A baixa representatividade do método de observação é preocupante. Os estudos em CO parecem não analisar o que as pessoas fazem e sim o que dizem que fazem. Isto pode ser explicado pelo alto custo de usar a observação e pela resistência das organizações em abrirem espaço para que ele seja usado nelas (Mourão *et al.*, 2016). Além disto, é um método de alta

precisão, baixa generalidade, por ser usado em amostras pequenas, e baixo realismo, considerando a interferência do pesquisador. Todos os problemas listados se aplicam aos métodos clássicos de observação, mas não quando ela é feita usando comportamentos em ambientes virtuais ou recursos tecnológicos para registrar os comportamentos. As novas tecnologias de informação e comunicação permitem medir uma imensa quantidade de comportamentos sem que haja qualquer interferência.

Os comportamentos gravados por meio de tecnologia trazem dois novos desafios, que são o tratamento e a análise dos dados (Guzzo *et al.*, 2015). Agregando o arcabouço teórico que a POT já tem, às soluções que já são usadas em áreas como a computação, a economia e a biologia, será possível conseguir dados mais precisos e fazer melhores generalizações, provavelmente mais próximas do que acontece durante o fazer humano no mundo do trabalho e das organizações. As novidades trazidas pelas inovações tecnológicas relacionadas a comportamentos são enquadradas em uma de duas categorias: armazenamento de informações e análise de dados.

Os comportamentos armazenados em banco de dados podem vir de meios virtuais (e.g., acesso a *sites*) ou não (e.g., gravação pela câmera do computador). Algumas informações armazenadas são muito mais complexas do que as já citadas e vão além da percepção consciente humana (e.g., dilatação da pupila, alterações galvânicas na pele, temperatura corporal e micro expressões faciais). Quanto à análise de dados, os benefícios do *big data* podem ser advindos da grande quantidade de dados, do grande poder estatístico ou de novas técnicas. Grandes bancos de dados facilitam a descoberta de padrões que não são aparentes em amostras menores. Uma aplicação desse benefício é a identificação de traços de personalidade a partir de comportamentos em redes sociais (Youyou, Kosinski & Stillwell, 2015).

Entre as novas técnicas de análise de dados estão a mineração textual e o mapeamento de redes de relacionamento, que permitem que sejam tiradas conclusões completamente novas. A aplicação desses benefícios para a POT ainda está em fase inicial. Já há alguns exemplos disponíveis. Por exemplo, a mineração de dados foi usada para melhorar a eficácia da seleção de pessoal, por Aghaee e Aghaee (2016) e Azar, Sebt, Ahmadi e Rajaeian (2013). Um algoritmo que identifica informações relevantes em currículos e em vagas e liga candidatos a vagas foi concebido por Paoletti, Martinez-Gil e Schewe (2015). Outro algoritmo, que prevê a nota de avaliação de desempenho dos funcionários, o que permitiria à área de gestão de pessoas fazer intervenções para melhorar o desempenho, foi desenvolvido por Kirimi e Moturi (2016). Um *software* chamado CORPNET, que analisa relacionamentos formais e informais para construir redes de relacionamento e identificar estruturas de poder e organizacionais foi criado por Liu e Moskivna (2016). Movimento de olhos, anotações, *logs* e reconhecimento facial para identificar processos autorregulatórios, durante a aprendizagem, foram utilizados por Azevedo *et al.*, (2013).

Há muito a explorar no uso dos métodos de *big data* em POT. Espera-se que transformem completamente a forma como informações são levantadas e decisões são tomadas nas organizações. Essa forma de pensar e agir poderá afetar todos os subsistemas de POT.

CAPÍTULO 2 – ESTILOS DE APRENDIZAGEM AUTODECLARADOS

Considerando as questões metodológicas de CO apresentadas no Capítulo 1, buscou-se um tema onde houvesse a possibilidade de relacionar a observação do comportamento, por meio de *machine learning*, ao uso de escalas. O tema escolhido foi o de estilos de aprendizagem, que foram analisados em sua forma tradicional, por meio de escalas (autodeclarados) e de uma forma inovadora, por meio de comportamentos em AVA (observados). O Capítulo 2 focou nas pesquisas em AVAs e analisou os conceitos de aprendizagem no trabalho e estilos de aprendizagem. A partir dessa análise, o objetivo do capítulo é criar uma escala de estilos de aprendizagem, com bons indícios de validade estatística, adequados para serem observados em AVA.

2.1 Introdução

2.1.1 Ambientes Virtuais de Aprendizagem

Analisar o “fazer” em organizações pode ser realizado de várias formas. Atualmente as organizações possuem uma grande massa de dados (*big data*) disponível, embora elas mesmas raramente consigam analisar. Uma das áreas que busca analisar *big data* é chamada de *analytics*, que é uma disciplina na intersecção entre ciência da computação e engenharia que usa métodos quantitativos para organizar, analisar e tirar conclusões do grande volume de dados gerado (Mortenson, Doherty & Robinson, 2015). Em conexão com a parte estratégica das organizações, esta disciplina pode ser entendida como um desdobramento de *Business Intelligence* (BI) (Davenport & Harris, 2007).

A aplicação de *analytics* para a área de recursos humanos (RH) é chamada de *HR analytics* ou *people analytics*. Ganhou grande popularidade nos últimos anos, sendo considerada não só uma tendência como uma realidade (Deloitte, 2017). Para os padrões de *big data*, a quantidade de informações atualmente disponíveis sobre funcionários é

considerada pequena. Contudo, em comparação com as informações geralmente disponíveis para acadêmicos de ciências sociais, a quantidade é imensa. Informações não estruturadas, como e-mails, comunicação por redes sociais, pesquisas na *web*, imagens, localização, entre outros, e informações tradicionais de recursos humanos, como posições ocupadas na empresa, notas de avaliação de desempenho, entre outros, são conjuntamente usadas em *HR analytics*. Para isso, é necessária a integração de dados de várias fontes e a compreensão deles no contexto em que estão inseridos. No entanto, os profissionais de RH frequentemente não teriam as habilidades, o conhecimento e os *insights* necessários para fazer as perguntas corretas usando os dados que têm à disposição, além de terem como premissa que as pessoas não podem ser reduzidas a métricas (Angrave, Charlwood, Kirkpatrick, Lawrence & Stuart, 2016). Por outro lado, os profissionais de ciências exatas, como ciência da computação e engenharia, não têm o conhecimento necessário sobre comportamento humano para identificar as informações efetivamente relevantes neste tipo de banco de dados. Isto é, existe um abismo entre aqueles e estes, ou talvez entre teoria e método.

Um tipo de ambiente presente em organizações são os AVAs, onde funcionários realizam treinamentos a distância, mas que também estão presentes no contexto da educação, por exemplo, em instituições de ensino com a modalidade a distância. Existem muitos ambientes que se encaixam nessa categoria e eles podem tanto ser soluções prontas e disponíveis no mercado, quanto soluções personalizadas para uma organização ou objetivo específico. Dos AVAs disponíveis no mercado, os mais utilizados são o *Moodle* e o *Blackboard*. Eles gravam informações como acesso a materiais e duração de sessões de acesso ao ambiente, pelos participantes, e avaliações de aprendizagem atribuídas a tais participantes. Ambientes personalizados podem gravar ainda mais informações, como o movimento do *mouse*, o tempo de duração da visualização de cada parte de um documento, ou *links* externos acessados. Além disso, esses ambientes também podem ser montados de

forma a interferirem intencionalmente no comportamento do usuário como, por exemplo, uma escolha forçada entre dois recursos. Todas essas informações, apesar de estarem em forma bruta, podem ser diretamente associadas a comportamentos do usuário no ambiente virtual e podem ser interpretadas. São chamadas de *footprints*, ou “pegadas” ou rastros virtuais. Sem tratamento, não são úteis. No entanto, *footprints* podem ser transformadas em indicadores de comportamentos observáveis. Um exemplo de indicador é a diferença de tempo entre a entrega de uma tarefa de treinamento e seu prazo final definido pelo curso disponibilizado em AVA.

A presente Dissertação se propõe a analisar um AVA usando ferramentas de POT e ciência da computação, unindo o conhecimento de comportamento humano com a manipulação e análise de *big data*. Isso será feito analisando os estilos de aprendizagem dos estudantes por medidas tradicionais (questionário) e inovadoras (indicadores de comportamento baseados em *footprints*) e comparando-as. A forma principal escolhida para a análise foi a detecção automática de estilos de aprendizagem, tema que será analisado a fundo por meio de revisão de literatura. Para fazer a detecção automática dos estilos, foi construída uma nova escala de estilos de aprendizagem autodeclarados. A seguir, será descrita a elaboração dessa escala.

2.1.2 Aprendizagem no trabalho

A aprendizagem faz referência a mudanças no comportamento do indivíduo não resultantes unicamente de maturação, mas da interação dele com o contexto (Abbad & Borges-Andrade, 2004). A aprendizagem no trabalho é aquela que ocorre no contexto do trabalho e pode ser definida como um processo individual de aquisição de conhecimentos,

habilidades e atitudes, que resulta em mudança de comportamento, decorrente de interação social situado em locais onde ocorre trabalho (Cairns & Malloch, 2011).

Esse fenômeno é de grande importância para organizações, pois o contexto de trabalho está cada vez mais dinâmico. A globalização e o surgimento de novas tecnologias tornaram o trabalho cada vez mais complexo, fluido, invisível, com maior incerteza e interligado (Borges-Andrade, Abbad & Mourão, 2006). Dessa forma, o trabalhador necessita mais qualificação para realizar seu trabalho. Essa qualificação pode ser advinda da educação, de treinamentos ou de eventos não-estruturados que ocorrem no trabalho. Aprendizagem informal é a aprendizagem espontânea, que ocorre independentemente da iniciativa deliberada da organização, sem estruturação prévia de condições de ensino-aprendizagem.

Aprendizagem formal ocorre quando a aquisição é promovida por iniciativa da organização de trabalho, que intencionalmente estrutura um contexto para que seus membros aprendam (Abbad & Borges-Andrade, 2004).

Os modelos de instrução têm sido amplamente utilizados para a aprendizagem formal. Três gerações de modelos instrucionais foram apontados por Kraiger (2008). A primeira geração consiste em analisar a tarefa a ser realizada e as necessidades de desempenho, desenhar o conteúdo, selecionar os métodos de entrega e avaliar a efetividade. O *instructional systems design* (ISD) oferece uma abordagem sistemática de identificação de uma lacuna de aprendizagem, de uma proposição de ensino-aprendizagem que supra a necessidade e da avaliação dessa solução (Gagné & Briggs, 1974). Essa abordagem é centrada no instrutor, que tem o papel de ensinar e coloca o estudante como passivo nesse processo. A segunda geração de modelos instrucionais considera que os objetivos são individualmente construídos e que o aprendiz os constrói com base no contexto em que o conhecimento deve ser aplicado. O aprendiz tem um envolvimento ativo selecionando, agindo sobre e organizando o novo conhecimento. Esta segunda geração sugere que os modelos mentais do aprendiz mediam a

relação entre a aquisição do conhecimento e o desempenho. Essa abordagem não funcionaria para treinamentos *online* (Kraiger, 2008). O aprendiz não possui bom conhecimento sobre suas necessidades de aprendizagem e o contexto, além de ter dificuldade de organizar seu próprio processo de aprendizagem. A terceira geração tem como base a abordagem social construtivista. Essa abordagem considera que as interações sociais são fundamentais para que a aprendizagem ocorra. O processo social de desenvolver um entendimento compartilhado por meio da interação seria a forma ‘natural’ das pessoas aprenderem (Hiltz, 1994). Essa terceira geração considera três tipos de interação: do aprendiz com o conteúdo, do aprendiz com o instrutor e do aprendiz com outros aprendizes. Para que cursos *online* sejam bem-sucedidos em gerar aprendizagem, eles deveriam tanto considerar o aprendiz como ativo no processo de aprendizagem quanto fornecer oportunidades de interação (Kraiger, 2008).

Considerando o aprendiz em cursos onde ele tem um papel ativo, variáveis individuais, variáveis de desenho do treinamento e o suporte dado ao aprendiz afetam as decisões tomadas por eles durante o treinamento (Noe, Clarke & Klein, 2014). Dessa forma, essas variáveis devem ser entendidas ao criar um treinamento onde a aprendizagem é controlada pelo aprendiz para potencializar seus resultados.

Uma das maiores dificuldades na área de aprendizagem formal no trabalho é mensurar de forma efetiva os resultados de um evento instrucional, as variáveis que interferiram no processo e como elas se influenciam. As pesquisas da área de treinamento e desenvolvimento nos últimos 100 anos foram analisadas em uma revisão publicada por Bell, Tannenbaum, Ford, Noe & Kraiger (2017). A partir da análise, eles categorizaram as variáveis analisadas em quatro grupos: (i) critérios de avaliação de aprendizagem; (ii) características do aprendiz; (iii) desenho do treinamento e entrega; (iv) contexto do treinamento.

Quanto aos critérios de avaliação, como não é possível ver a aprendizagem diretamente, apenas seus efeitos, o entendimento do que é aprendizagem foi expandido.

Inicialmente, usava-se como base a abordagem behaviorista, que considerava a aprendizagem como mudança de comportamento e, subsequentemente, passou-se a usar a abordagem cognitivista, que entende que pode haver aprendizagem sem mudança de comportamento. A partir da abordagem cognitivista, foram desenvolvidas as hierarquias de aprendizagem, que dividiram a aprendizagem em relação ao domínio e à complexidade. Dessa forma, foi possível analisar de forma inter-relacionada vários construtos de aprendizagem. Passou-se a analisar separadamente a aquisição e a transferência da aprendizagem. Além disso, a área também passou a estudar os efeitos da transferência da aprendizagem na equipe e na organização, ampliando os tipos de resultados.

Outra categoria é a das características dos aprendizes, que podem ser divididas em cinco grupos: (i) capacidades; (ii) traços de personalidade; (iii) construtos motivacionais; (iv) valores e interesses; (v) atitudes e emoções. O foco inicial era prever o sucesso no treinamento para a alocação eficiente de recursos. Nos anos 1990, três novas áreas de pesquisa surgiram: (i) relação entre características do aprendiz e busca por desenvolvimento; (ii) efeito do treinamento na socialização dos aprendizes; (iii) relação entre autoeficácia, conscienciosidade, orientação para objetivos, clima e cultura e a interação com motivação para aprender, desenvolvimento de objetivos de aprendizagem, análise de tarefas, atividades metacognitivas e aprendizagem e transferência de treinamento. Dessa forma, passou-se a entender que há uma relação complexa entre variáveis individuais e outras variáveis ligadas ao treinamento e ao trabalho.

As outras duas categorias identificadas pelos autores foram o treinamento e a entrega, onde eles identificaram a mudança do estudante de passivo para ativo, a intensificação do uso da tecnologia e o uso de treinamento para melhorar o funcionamento de times. Com essas mudanças, as variáveis de contexto passaram a ter um papel importante na análise de treinamentos. Variáveis de suporte tanto antes quanto após o treinamento foram amplamente

estudadas. O suporte apresentou uma relação positiva tanto com a motivação para aprender quanto para transferir.

Cada vez mais investimentos são feitos pelas organizações em treinamentos, uma vez que a aprendizagem no trabalho é um diferencial competitivo e precisa estar conectada com os objetivos estratégicos da organização (Abbad, Zerbini & Souza, 2010). A aprendizagem informal ganhou espaço tanto nos estudos acadêmicos brasileiros e internacionais (Coelho Jr. & Borges-Andrade, 2011; Noe, Clarke & Klein, 2014) quanto nas organizações. Mesmo assim, os treinamentos formais continuam recebendo a maior parte do investimento.

Considerando o contexto do trabalho, o trabalhador busca cada vez mais a aprendizagem contínua. No contexto das organizações, isso se reflete na aprendizagem se tornando cada vez mais controlada pelo aprendiz, assim como afetada socialmente e reconhecida como uma ocorrência natural do trabalho (Noe, Clarke & Klein, 2014). No caso dos treinamentos formais, o que possibilitou os avanços na centralização do processo de aprendizagem no indivíduo foram as novas tecnologias.

Ferramentas como AVAs e aplicativos para *smartphones* e *tablets* permitiram que o ensino passasse a ser assíncrono, personalizado e realizado em qualquer lugar. O método tradicional, de ensino presencial, é caro e impossibilita a realização por pessoas com grande dispersão geográfica (Noe, Clarke & Klein, 2014). Por outro lado, o custo do ensino a distância é concentrado na elaboração, sendo diluído quando oferecido a grandes grupos. Além disso, em muitos casos, os cursos em EAD são elaborados de forma que não haja nem a necessidade de um instrutor. As vantagens do EAD fizeram com que a procura por essa modalidade aumentasse nos últimos anos, uma vez que ela é mais adequada tanto em relação à realidade das empresas quanto ao ensino de adultos. Contudo, aquele custo pode ficar bastante elevado, quando ocorre elevada evasão, o que é frequente em EAD. Esta evasão

muitas vezes é associada a determinados estilos de aprendizagem, que poderiam ser automaticamente detectados no AVA.

2.1.3 Estilos de aprendizagem

Haveria quase tantas definições de estilos de aprendizagem quanto teóricos, apontou De Bello (1990). Houve várias tentativas de agrupar e integrar a literatura da área. A mais utilizada foi a taxonomia de Coffield, Moseley, Hall e Ecclestone (2004). Neste estudo, foram analisados mais de 800 artigos e as teorias foram posicionadas em um contínuo de cinco famílias: (i) de base constituinte; (ii) estruturas cognitivas; (iii) traços estáveis de personalidade; (iv) preferências de aprendizagem flexíveis-estáveis; e (v) abordagens, estratégias, orientações e concepções de aprendizagem.

A família de estilos de base constituinte considera que a combinação de fatores genéticos e de desenvolvimento leva o indivíduo a desenvolver um determinado estilo, que dificilmente muda. Os principais modelos dessa família são os de Dunn e Dunn (Dunn, Dunn & Price, 1981), de Gregorc (Gregorc, 1979).

A família das estruturas cognitivas entende que os estilos refletem recursos profundamente arraigados do aparato cognitivo, como conjuntos de capacidades. Os principais modelos dessa família são os de Gardner (Gardner, 1983) e Kagan (Kagan & Kogan, 1970).

A família dos traços estáveis de personalidade é composta por modelos que têm como base Carl Jung e que relacionam a personalidade com a forma de aprender. Os principais modelos dessa linha são os de Apter (Apter, Mallows & Williams, 1998), Jackson (Jackson, 2002) e Myers-Briggs (Myers, McCaulley, Quenk & Hammer, 1998).

A família das preferências de aprendizagem flexíveis-estáveis é composta por modelos que entendem os estilos como preferências relativamente estáveis que podem se expressar de

diferentes formas de acordo com o contexto. É dessa categoria que fazem parte a maior parte dos modelos usados para a detecção automática. Os principais modelos da família são o de Allinson e Hayes (Allinson & Hayes, 1996), de Honey e Mumford (Honey & Mumford, 1992), de Kolb (Kolb, 1976), de Felder e Silverman (Felder, 1996) e de Berings (Berings, Poell & Simons, 2005), que surgiu depois da criação da taxonomia de Coffield.

Por último, a família que estuda abordagens, estratégias, orientações e concepções considera que o contexto e as experiências anteriores influenciam os estilos, o que faz com que eles sejam flexíveis. Os principais modelos dessa família são os de Entwistle (Entwistle, 1981), Sternberg (Sternberg, 1999) e Vermunt (Vermunt, 1994).

Na área de treinamento e desenvolvimento, os estilos são usados como uma variável que pode potencializar resultados de indivíduos, grupos e de toda a organização (Sternberg, Zhang & Rayner, 2011). A aplicação mais usada dos estilos em ambientes virtuais é a individualização do ensino, de onde fazem parte os sistemas adaptativos, que foram explicados anteriormente. Esse uso recebe muitas críticas porque muitas vezes são usados instrumentos com baixa validade psicométrica e com problemas conceituais e empíricos (Salles, 2007). Profissionais de áreas que não são ligadas à educação, principalmente ciência da computação e engenharia, usam os estilos de aprendizagem sem terem um conhecimento aprofundado da teoria, o que pode levar a erros em sua aplicação.

Uma revisão de literatura de estilos de aprendizagem, analisando as produções nacional e internacional, foi feita por Jesus (2014). Ela encontrou apenas 14 artigos empíricos nacionais usando a variável estilo de aprendizagem e, destes, só 3 se relacionavam ao contexto de trabalho. Os outros onze foram trabalhos desenvolvidos no contexto educacional que visam descrever o perfil de estudantes da educação básica, média e superior, indicar melhorias curriculares para atender a perfis diferentes de aprendizes e prever o rendimento acadêmico. Além dos artigos, foram identificadas 48 dissertações e 12 teses. Com aplicações

no contexto de trabalho, houve apenas a dissertação de Salles (2007), no Brasil. Sobre a pesquisa internacional, tipicamente os modelos e instrumentos são usados para categorizar os alunos, atribuindo a eles um perfil individual de aprendizagem. De forma simples, entende-se que os aprendizes que estão em uma determinada categoria aprendem melhor de uma determinada forma.

Não há evidências suficientes para afirmar que o ensino adaptado aos estilos de aprendizagem é melhor do que o modelo tradicional (Pashler, McDaniel, Roher & Bjork, 2009). A crítica recai sobre a categorização dos indivíduos, a falta de qualidade dos estudos empíricos e o apelo comercial. O principal problema apontado é que a adoção do ensino adaptado aos estilos pode levar à particularização excessiva da instrução. Os esforços deveriam ser concentrados em desenvolver medidas válidas e desenhar estudos que possam demonstrar o efeito dessa variável na aprendizagem, apontam Pashler *et al.* (2009). Uma outra revisão sobre estilos de aprendizagem e estratégias de ensino indica a abundância e a fragilidade dos modelos teóricos e dos instrumentos de medida (Wilson, 2012).

Salles (2007) analisou 57 medidas referentes a 12 modelos de estilos de aprendizagem aplicados no contexto do trabalho. As medidas foram analisadas quanto à validade, sendo ela de construto, preditiva e de conteúdo, e à confiabilidade. Os modelos analisados estão listados na Tabela 2. Apenas duas das medidas analisadas foram consideradas válidas e confiáveis. Os problemas encontrados foram tanto das medidas quanto da possibilidade de aplicação da teoria em um determinado contexto. Dessa forma, entende-se que é necessário analisar a teoria que está sendo utilizada para os estilos quanto a sua aplicabilidade no contexto de pesquisa e os instrumentos disponíveis quanto à capacidade deles de medirem os construtos da teoria.

Tabela 2
Instrumentos de estilos de aprendizagem utilizados em contexto de trabalho

Autor(es)	Medida	Ano de publicação
Myers-Briggs	Myers-Briggs Type Indicator (MBTI)	1962
Witkin	Group Embedded Figures Test (GEFT)	1971
Kolb	Learning Style Inventory (LSI)	1971
Kirton	Kirton Adaptation-Innovation Inventory (KAI)	1976
Hunt, Butler, Noy e Rosser	Paragraph Completion Method (PCM)	1978
Price, Dunn e Dunn	Productivity Environmental Preference Survey (PEPS)	1979
Rezler e Rezmovic	Learning Preference Inventory (LPI)	1981
Gregorc	Gregorc Style Delineator (GSD)	1982
Honey e Mumford	Learning Styles Questionnaire (LSQ)	1982
Herrman	Herrman Brain Dominance Instrument (HBDI)	1982
Epstein e Meier	Constructive Thinking Inventory (CTI)	1989
Allinson e Hayes	Cognitive Style Index (CSI)	1996

Fonte: Salles, T. (2007). Estilos de aprendizagem no trabalho: análise e construção de medidas. Dissertação de Mestrado. Universidade de Brasília. Brasília, DF.

Um exemplo de problema de aplicabilidade dos modelos é a aplicação do modelo Felder-Silverman para a detecção automática de estilos de aprendizagem. Quanto à teoria, os indivíduos são classificados em quatro contínuos e não é prevista a categorização dentro deles. Além disso, as quatro dimensões dos estilos se referem a preferências, mas a aprendizagem é potencializada quando o estudante aprende tanto de sua forma preferencial quanto da que ele não prefere. Por exemplo, estudantes que têm um estilo auditivo aprendem melhor quando recebem estímulos auditivos e verbais. Na teoria, também é salientada a importância de o estudante ser apresentado a formas de aprendizagem que ele não prefere para que ele desenvolva a habilidade de aprender de várias formas. Quanto ao instrumento, o *Index of Learning Style (ILS)*, a consistência interna de seus fatores é baixa, com alfas de Cronbach que variam de 0,41 a 0,76 (Felder & Spurlin, 2005).

Além disso, é necessário considerar que a modalidade a distância apresenta características e necessidades diferentes do ensino presencial e que os estilos de aprendizagem podem ter uma relação diferente com a instrução nessa modalidade. Algumas características

apontadas como fundamentais para melhorar persistência em ensino a distância são os comportamentos de autorregulação (e.g., gerenciamento do tempo) e a comunicação com o tutor (Hart, 2012). As principais variáveis apontadas como fundamentais para o sucesso do estudante em ensino a distância são a inteligência emocional, expressa por auto percepção de necessidades e manejo de sentimentos, habilidades de autorregulação, disciplina, gestão do tempo, organização, planejamento, auto avaliação, estilos de aprendizagem reflexivos e visuais e locus de controle interno (Kauffman, 2015). Os estilos de aprendizagem afetam de forma diferente os estudantes em ensino nos contextos presencial e a distância e características como a autorregulação e a interação têm um papel central nesse segundo contexto.

Um dos modelos que considera a regulação da aprendizagem como parte dos estilos é o de Vermunt (Vermunt, 1996), da família de abordagens, estratégias, orientações e concepções para a aprendizagem (Coffield *et al.*, 2004). Denominado de Modelo de Regulação dos Processos de Aprendizagem, é composto por dimensões, que se relacionam de acordo com a Figura 3: (i) estratégias de processamento cognitivo; (ii) estratégias de regulação da aprendizagem; (iii) modelo mental de aprendizagem; e (iv) orientações para a aprendizagem. Os estilos de aprendizagem são uma visão geral de estratégias que o estudante geralmente emprega, a orientação que ele geralmente tem para a aprendizagem e os modelos mentais que ele geralmente utiliza. As estratégias de processamento cognitivo são as atividades mentais, como identificar quais são as partes principais de um conteúdo, e externas, como marcar nos textos de estudo as partes identificadas, que o indivíduo usa para processar o conteúdo. As estratégias regulatórias são as atividades decorrentes do monitoramento cognitivo e afetivo da aprendizagem como, por exemplo, buscar textos além dos que foram indicados na disciplina. O modelo mental de aprendizagem é um sistema coerente de concepções do indivíduo sobre a aprendizagem e sobre o conhecimento como, por exemplo,

que os conhecimentos de diferentes disciplinas são interligados. Por último, a orientação para aprendizagem tem como base objetivos que os estudantes têm em relação ao evento instrucional como, por exemplo, estudantes que têm como objetivo apenas se formarem em uma graduação (Vermunt, 1996). A forma como os estudantes processam o conteúdo é determinada principalmente pelas estratégias regulatórias que ele adota. Os modelos mentais e a orientação para a aprendizagem também afetam o processamento, mas a maior parte dessa influência é mediada pelas estratégias regulatórias (Figura 5).

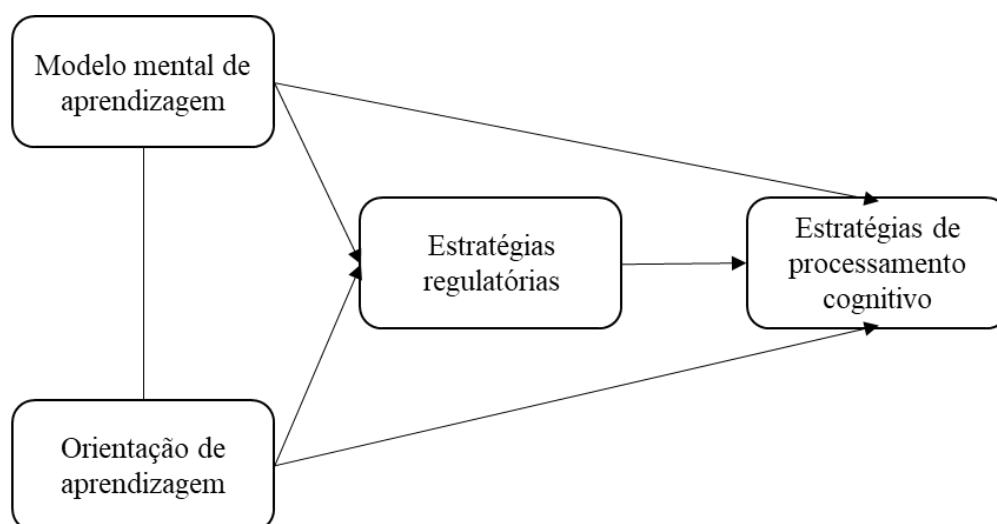


Figura 5. Modelo de Regulação dos Processos de Aprendizagem

Fonte: Vermunt, J. D. (1998). The regulation of constructive learning processes. *British Journal of Educational Psychology*, 68, 149-171.

Outro modelo da mesma família é o de Entwistle (Entwistle, 1998). Nesse modelo, as intenções e objetivos dos estudantes determinam sua orientação educacional. Ela pode se relacionar a motivações extrínsecas e intrínsecas e essas motivações podem mudar de acordo com o contexto. Os estudantes também têm concepções de aprendizagem. Alguns estudantes entendem a aprendizagem como aumentar o conhecimento ou lembrar fatos, que é considerada uma concepção sem sofisticação, enquanto outros entendem que a aprendizagem depende da abstração do sentido e que o que se entende da realidade é baseado em interpretações, que é considerada uma concepção sofisticada (Entwistle, 1990). As

orientações e concepções se expressam na abordagem que o estudante adota para a aprendizagem. Essa abordagem pode ser superficial, estratégica ou profunda (Entwistle, 1998). Os conceitos e instrumentos das teorias de Entwistle e Vermunt são muito relacionados (Entwistle & Peterson, 2004). Um dos instrumentos utilizados é o ASSIST, que contém três escalas (concepções de aprendizagem, abordagens para a aprendizagem e preferência de ensino). A escala de abordagem para a aprendizagem possui três subescalas (abordagem profunda, abordagem estratégica e abordagem superficial) e cada uma delas possui quatro subescalas (Entwistle, McCune & Tait, 2013).

O Modelo de Regulação dos Processos de Aprendizagem foi usado por Jesus (2014) para construir uma medida de estilos de aprendizagem em ensino a distância. Essa medida tem boas evidências de validade e é adequada para ser usada para cursos a distância. A escala foi denominada Estilos Preferenciais de Aprendizagem a Distância e é composta por três fatores: (i) autorregulação de aprendizagem; (ii) regulação externa; e (iii) interação. O mesmo modelo foi usado por Moraes e Zerbini (2018) e Moraes (2016) na construção de uma escala de estilos de aprendizagem em EAD. A versão final da escala conta com 8 itens, organizados em 3 fatores: (i) ambiente de estudo; (ii) regulação externa; e (iii) interação interpessoal.

Todos os instrumentos citados nessa revisão referem-se a estilos de aprendizagem autodeclarados. Como todas as medidas são questionários, o que está sendo analisado é o “dizer” e não o “fazer”. Além disso, em todas as medidas os participantes respondem sobre eles mesmos, então, o que declaram sobre eles mesmos é a única fonte de informação. Tendo em vista o objetivo específico 3 dessa Dissertação, que é detectar os estilos de aprendizagem autodeclarados por meio de comportamentos em AVA, e que, para atingir esse objetivo, foram analisados vários cursos simultaneamente, foram escolhidos como teoria de estilo, Vermunt (1998), e como instrumento-base, Jesus (2014). Essa escolha deve-se a usarem estilos que são relativamente estáveis e que estão pouco ligados ao conteúdo do curso. Além

disso, entendem que existem preferências de aprendizagem que influenciam os comportamentos no curso. Ou seja, os comportamentos podem variar de acordo com o curso, mas os indivíduos têm uma tendência que varia pouco com o tempo. Nas seções subsequentes foram apresentadas as adaptações feitas no questionário, o método e as validações exploratória e confirmatória.

2.2 Método

2.2.1 Construção da escala

A escala de Jesus (2014) foi alterada com dois objetivos: adequá-la à observação dos comportamentos no AVA e abarcar o máximo possível de comportamentos que podem ser identificados em um banco de dados de um Moodle sem qualquer *add-on*. Como os estilos serão posteriormente identificados por um algoritmo que tem como base comportamentos em AVA, a aproximação teórica do questionário aos comportamentos pode diminuir o erro. Por isso, outra adaptação que foi feita foi que o respondente indicasse seu padrão comportamental e não sua preferência. Uma preferência é uma atitude em relação a um objeto que tipicamente se expressa em uma tomada de decisão (Lichtenstein & Slovic, 2006). Apesar da preferência estar associada a uma atitude que afeta diretamente a frequência de um comportamento, ela não é a própria frequência do comportamento. Por isso, o questionário foi analisado e modificado, para que se transformasse em uma medida de frequência.

Foram investigados os cursos a distância de um órgão da administração pública federal para o estudo dos estilos de aprendizagem e o atingimento do objetivo geral da Dissertação, assim como dos quatro objetivos específicos. A plataforma utilizada pela instituição é o Moodle e ela conta com 305 cursos a distância com cargas horárias de 10 a 60 horas, que são disponibilizados para os servidores do órgão e para toda a sociedade. As informações de todos os cursos, incluindo os estudantes que os realizaram e as informações

armazenadas no sistema até novembro de 2017, foram disponibilizadas para a realização deste trabalho. Ao todo, a base de dados contém 29.173 estudantes cadastrados.

Inicialmente as informações dos bancos de dados foram analisadas para identificar quais comportamentos poderiam ser observados no AVA. Foram mapeadas 211 tabelas de acordo com as informações presentes em cada uma. Uma estrutura relacional do banco de dados (Apêndice C) foi fornecida aos pesquisadores. As informações verificadas foram categorizadas em tarefas, notas, interação no curso, informações do curso, comportamentos dos usuários em cada curso e informações dos usuários. As informações de comportamentos dos usuários continham dados de quando foram realizadas as atividades e, em alguns casos, da duração das mesmas.

Para a elaboração do questionário foram adotados procedimentos metodológicos descritos por Pasquali (2010). São sugeridas as etapas: (i) definição do sistema psicológico de interesse; (ii) identificação das propriedades do sistema psicológico; (iii) caracterização da dimensionalidade; (iv) conceituação por definição constitutiva e operacional; (v) operacionalização; e (vi) análise teórica dos itens. O presente estudo busca analisar o processo de aprendizagem induzida com foco em comportamentos. A variável de interesse dessa análise são os estilos de aprendizagem. Estilo de aprendizagem representa um comportamento preferido para aprender em treinamento a distância (Jesus, 2014). De forma mais ampla, o estilo de aprendizagem é um construto multideterminado composto por preferência de processamento de informação, por estratégias de regulação, por concepções sobre a aprendizagem e pelos interesses pessoais que orientam a aprendizagem (Vermunt, 1996). Aqui, ele é entendido como um padrão comportamental frequente e expresso em diversos contextos de AVA. Apesar da frequência dos comportamentos variar com o contexto, entende-se que sua frequência relativa se mantém, até certo ponto, constante.

Para a construção da Escala de Estilos de Aprendizagem em Ensino a Distância com base em Comportamentos (EEA-Ead_C), foi usada como base a escala de estilos preferenciais de aprendizagem a distância (Jesus, 2014). Sua estrutura tem três fatores que foram identificados pela autora, com base em evidências de validação psicométrica: (i) autorregulação de aprendizagem; (ii) regulação externa; e (iii) interação. Após a análise preliminar da estrutura relacional do banco de dados, foram identificados os comportamentos que poderiam ser identificados no AVA. Para aproximar os itens, essa escala de estilos preferenciais de aprendizagem a distância, de comportamentos observáveis, eles foram comparados com esses comportamentos, que seriam expressos no AVA na forma de indicadores. Por exemplo, não é possível observar no banco de dados se o estudante buscou materiais além do curso para ajudar em seu estudo, no entanto, é possível observar se ele acessou todos os materiais que estavam disponíveis. Dessa forma, ao comparar os itens com o banco, eles foram alterados para aumentar a quantidade de itens da escala com indicadores comportamentais correspondentes no AVA. Além disso, foram analisados outros instrumentos disponíveis, visando buscar mais itens que pudessem ser observados no banco e que estivessem de acordo com a definição teórica do constructo, nos relatos de Salles (2007) e de Tait, Entwistle e McCune (1998). Os itens de processamento cognitivo foram criados sem a preocupação com a referência em indicadores, uma vez que nenhum deles pode ser observado no banco de dados. Para a análise teórica do instrumento, foram agrupados os fatores de auto regulação e fatores externos de regulação em uma única categoria, a regulação. Essa categoria foi subdividida em: o estudo organizado, o gerenciamento do tempo e a potencialização da efetividade. O fator de interação também foi dividido em: aprendizagem vicariante e interação ativa (Quadro 1). A primeira versão construída continha 53 itens.

Quadro 1
Categorias da validação por juízes

Categorias	Definição
Regulação da aprendizagem <i>Estudo organizado</i> <i>Gerenciamento de tempo</i> <i>Potencializar efetividade</i>	Padrões gerais de comportamentos sistematizados para o atingimento de metas. <i>Comportamentos de sistematização do estudo do conteúdo.</i> <i>Comportamentos de controle consciente do momento ou do tempo gasto para realizar uma atividade.</i> <i>Comportamentos direcionados para melhorar a aprendizagem e para obter melhores notas.</i>
Processamento cognitivo	Atividades mentais de processamento de informação, planejamento e tomada de decisão que não podem ser observadas diretamente por comportamentos.
Interação <i>Aprendizagem vicariante</i> <i>Interação ativa</i>	Contato por meio de linguagem verbal com outras pessoas. <i>Observação do contato de terceiros com outras pessoas ou consigo, mas sem resposta do próprio sujeito.</i> <i>Contato do sujeito com outras pessoas.</i>

Para a validação por juízes, foram realizadas duas reuniões presenciais com oito estudantes de pós-graduação e profissionais das áreas de psicologia, pedagogia e administração. Na primeira reunião, foi apresentado o escopo do trabalho e foram apresentados os conceitos utilizados. Ao fim da reunião, cada um dos juízes respondeu a um instrumento onde deveriam indicar a categoria e a subcategoria que entendiam que cada um dos itens se encaixava. Na segunda reunião, foram apresentados os resultados da análise descritiva do pertencimento a cada categoria e subcategoria e os itens foram revisados um a um. Neste momento também foi apresentado aos juízes se o item era referente a um comportamento observável em AVA ou não (Apêndice A). Dessa forma, o objetivo foi representar satisfatoriamente as categorias e subcategorias definidas com o máximo possível de comportamentos observáveis em AVA, mas com o mínimo possível de redundância. Cada item foi revisado quanto a sua linguagem, relevância e pertencimento ao fator.

Após o procedimento, as sugestões de alteração, exclusão e inclusão de itens foram analisadas e o instrumento final contou com 31 itens. Destes, 14 foram de regulação, 9 de interação e 8 de processamento cognitivo.

O questionário foi construído de forma *online* na plataforma EaDuca que é uma plataforma onde podem ser criados *surveys*. Nela são cadastradas as questões, os enunciados e os participantes e, a partir dessas informações, são gerados *links* para a pesquisa. Os *links* dos questionários foram individualizados e enviados a todos os estudantes que tinham cadastro na plataforma (N = 29.173). A individualização dos *links* permitiu que as respostas fossem ligadas aos comportamentos no AVA. Dessa forma, os respondentes não precisaram informar seus dados pessoais, pois essas informações já estavam disponíveis e, por isso, não houve perda de respondentes. Para incentivar a resposta à pesquisa, foi fornecido um relatório simplificado do estilo de aprendizagem de cada respondente, com base na teoria utilizada (Apêndice D). A escala de resposta foi *Likert* de frequência, sendo composta por 5 pontos (1 – nunca; 2 – quase nunca; 3 – às vezes; 4 – quase sempre; e 5 – sempre). Foram contabilizadas 1.764 (6% do total de estudantes) respostas à pesquisa em um período de três meses, de outubro a dezembro de 2018.

2.2.2 Análises fatoriais exploratória e confirmatória

Para a análise fatorial exploratória, foi escolhida uma amostra aleatória de 310 respondentes para a análise exploratória de evidência de validade estatística. Com essa quantidade, há 10 respondentes para cada item da escala. Não houve casos omissos, pois o sistema não permitia que o respondente finalizasse a pesquisa sem responder a todos os itens. Foi realizada a análise de *outliers*, de normalidade univariada, análise de Principais Componentes e, por fim, a fatoração pelos Principais Eixos. As análises de dados foram realizadas usando o *software* SPSS 23.

Para a análise fatorial confirmatória, foi usada uma amostra da população de respondentes excluindo os casos usados na análise fatorial exploratória. Essa amostra aleatória foi de 310 respondentes. Com essa quantidade, considerando que foi testada a estrutura com 18 itens, há 17,2 respondentes para cada item da escala. Foi realizada a análise

de *outliers*, de normalidade univariada, análise de Principais Componentes a análise dos Principais Eixos e, por fim, os testes de ajuste do modelo. As análises de dados foram realizadas usando os *softwares* SPSS 23 e o AMOS 23.

2.3 Resultados

2.3.1 Análise fatorial exploratória

Inicialmente foram analisados os *outliers* univariados. A quantidade somada de casos extremos identificados para cada uma das variáveis foi superior a 20% da amostra, dessa forma, optou-se por não retirar qualquer caso extremo univariado. Os casos extremos multivariados foram identificados pela distância de Mahalanobis. Os sete casos identificados por esse método foram retirados da amostra, ficando a amostra com 303 respondentes. A seguir, foi feita a análise de evidências de normalidade. Apenas um dos 31 itens apresentou distribuição de acordo com a normalidade. Foram testadas transformações para melhorar a normalidade dos dados (raiz quadrada, logaritmo, razão, entre outros). Nenhuma delas alterou de forma significativa a distribuição dos dados e, por isso, eles foram analisados sem transformação. O Apêndice E contém as análises descritivas dos 31 itens, assim como a assimetria e a curtose.

O primeiro procedimento de análise fatorial utilizado foi o dos Componentes Principais. A amostra foi considerada fatorável, pois o KMO encontrado foi 0,87 (Tabela 3), que é considerado muito bom. Houve menos de 40% de correlações maiores do que 0,3. Todas as comunalidades encontradas foram maiores do que 0,3 e menores do que 0,8 (Tabela 4).

Tabela 3
Testes KMO e Bartlett da análise fatorial exploratória

Medida Kaiser-Meyer-Olkin de adequação de amostragem.		0,869
Teste de esfericidade de Bartlett	Aprox. Qui-quadrado	3385,833
	gl	465
	Sig.	0,000

Tabela 4
Comunalidades da análise fatorial exploratória sem retirada de itens

	Inicial	Extração
Logo antes das avaliações do curso, reviso os materiais de estudo.	1,000	0,633
Realizo todas as atividades do curso, mesmo as que não são pontuadas.	1,000	0,627
Realizo um curso de cada vez.	1,000	0,585
Acesso os materiais adicionais do curso.	1,000	0,665
Defino minhas próprias metas para estudar os materiais do curso.	1,000	0,655
Entrego com antecedência todas as tarefas do curso que têm prazo de entrega.	1,000	0,699
Leio todos os feedbacks que o tutor passa sobre as tarefas que realizei no curso.	1,000	0,676
Verifico minhas notas do curso assim que elas são lançadas.	1,000	0,332
Acesso os materiais do curso na ordem que são apresentados.	1,000	0,510
Acesso os materiais do curso várias vezes.	1,000	0,746
Busco materiais além dos que são passados no curso.	1,000	,744
Divido os materiais de curso de forma a criar um ritmo de estudo que sigo até o final do curso.	1,000	0,699
Realizo todos os exercícios do curso para praticar o que aprendi.	1,000	0,582
Realizo apenas as atividades do curso que são pontuadas.	1,000	0,668
Aprendo os conteúdos do curso a partir de insights (ideias que surgem repentinamente).	1,000	0,703
Aprendo melhor sozinho os conteúdos do curso.	1,000	0,703

Leio todas as perguntas, respostas e comentários que meus colegas e tutor postam no curso.	1,000	0,281
Leio todos os tópicos que são postados nos fóruns do curso.	1,000	0,496
Interajo com meus colegas de curso e com o tutor sobre assuntos que não estão diretamente ligados ao conteúdo.	1,000	0,557
Posto minha dúvida quando tenho uma relacionada ao conteúdo do curso.	1,000	0,609
Quando tenho uma dúvida, pergunto a meus colegas de curso.	1,000	0,559
Interajo com o tutor durante o curso.	1,000	0,475
Interajo com meus colegas de curso enquanto o realizo.	1,000	0,473
No curso, leio todas as respostas a postagens que faço.	1,000	0,522
Primeiro compreendo a estrutura geral do conteúdo do curso para depois estudá-lo em detalhes.	1,000	0,590
Ao estudar os conteúdos do curso, busco conhecimentos fundamentados em fatos.	1,000	0,473
Ao estudar os materiais do curso, foco em fatos e dados.	1,000	0,527
Uso minha intuição (entender as coisas no momento, sem necessidade de realizar raciocínios complexos) para aprender o conteúdo do curso.	1,000	0,571
Reflito sobre o conteúdo do curso para melhor compreendê-lo.	1,000	0,336
Reflito sobre as possibilidades de aplicação no meu contexto profissional do que aprendi no curso.	1,000	0,606
Ao estudar o material do curso, foco em ideias para aprender.	1,000	0,428

Uma vez que a amostra foi considerada fatorável, foi realizada a fatoração pelos Eixos Principais com rotação promax para identificar a estrutura fatorial do questionário. Pelo critério dos autovalores superiores a um, foram apontados sete fatores (Tabela 5) e pelo *scree plot* foram indicados quatro fatores (Figura 6). Foi realizada uma simulação de Monte Carlo (Damásio, 2012), que apontou quatro fatores.

Tabela 5

Variância explicada e autovalores para a análise de Principais Componentes da análise fatorial exploratória

Fator	Autovalores iniciais			Somos de extração de carregamentos ao quadrado			Somas de rotação de carregamentos ao quadrado
	Total	% de variância	% cumulativa	Total	% da variância	% cumulativo	Total
1	7,429	23,964	23,964	6,968	22,476	22,476	5,599
2	3,053	9,848	33,811	2,574	8,302	30,778	5,020
3	2,123	6,850	40,661	1,627	5,247	36,025	2,877
4	1,489	4,803	45,464	0,916	2,956	38,981	3,110
5	1,380	4,452	49,916	0,850	2,743	41,725	2,741
6	1,209	3,899	53,815	0,742	2,392	44,117	1,419
7	1,048	3,381	57,196	0,490	1,582	45,699	3,780
8	0,983	3,170	60,365	0,394	1,270	46,968	0,927
9	0,908	2,931	63,296				
10	0,878	2,833	66,129				

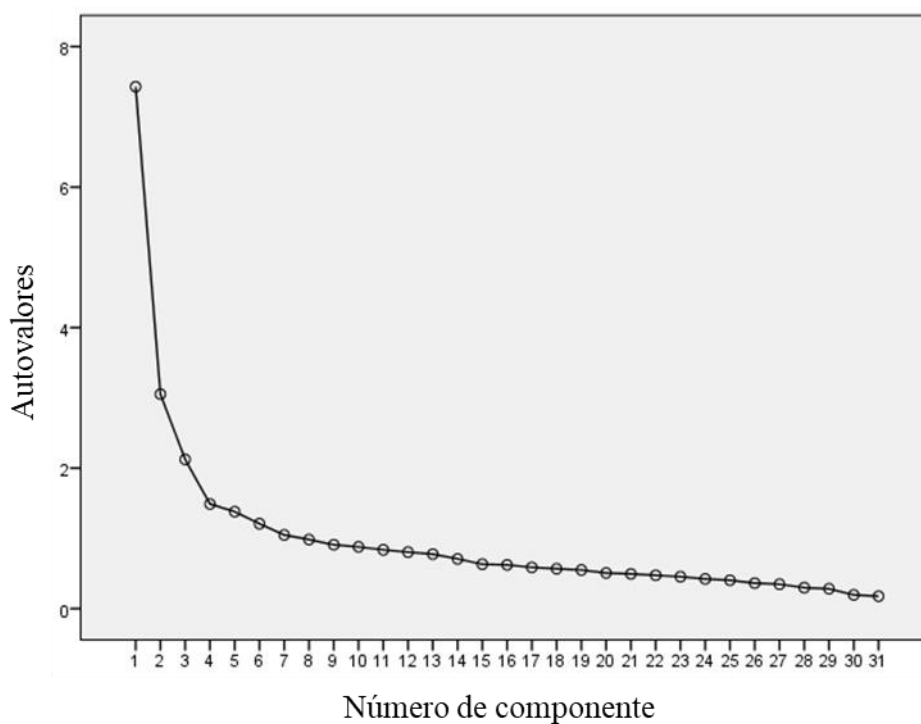


Figura 6. Scree plot da análise fatorial exploratória.

Foram testadas as estruturas de dois a oito fatores e todas com mais de três fatores se mostraram insustentáveis, uma vez que pelo menos um dos fatores apresentava quantidade insuficiente de itens. Sendo assim, escolheu-se a estrutura com 3 fatores. Esta apresentou

muitos itens com carga fatorial abaixo de 0,3 para todos os fatores e também algumas cargas fatoriais compartilhadas entre fatores (Tabela 6).

Tabela 6
Matriz de padrão da análise fatorial exploratória

	Fator		
	1	2	3
Logo antes das avaliações do curso, reviso os materiais de estudo.	0,360		0,391
Realizo todas as atividades do curso, mesmo as que não são pontuadas.	0,370		
Realizo um curso de cada vez.	0,483		
Acesso os materiais adicionais do curso.			0,452
Defino minhas próprias metas para estudar os materiais do curso.	0,367		0,409
Entrego com antecedência todas as tarefas do curso que têm prazo de entrega.			0,390
Leio todos os feedbacks que o tutor passa sobre as tarefas que realizei no curso.			0,473
Verifico minhas notas do curso assim que elas são lançadas.	0,440		
Acesso os materiais do curso na ordem que são apresentados.		0,395	
Acesso os materiais do curso várias vezes.		0,846	
Busco materiais além dos que são passados no curso.		0,886	
Divido os materiais de curso de forma a criar um ritmo de estudo que sigo até o final do curso.		0,829	
Realizo todos os exercícios do curso para praticar o que aprendi.		0,712	
Realizo apenas as atividades do curso que são pontuadas.		0,609	
Aprendo os conteúdos do curso a partir de insights (ideias que surgem repentinamente).	0,310	0,518	
Aprendo melhor sozinho os conteúdos do curso.		0,576	
Leio todas as perguntas, respostas e comentários que meus colegas e tutor postam no curso.		-0,423	
Leio todos os tópicos que são postados nos fóruns do curso.			0,405
Interajo com meus colegas de curso e com o tutor sobre assuntos que não estão diretamente ligados ao conteúdo.	0,683		
Posto minha dúvida quando tenho uma relacionada ao conteúdo do curso.	0,510		
Quando tenho uma dúvida, pergunto a meus colegas de curso.	0,536		
Interajo com o tutor durante o curso.	0,611		
Interajo com meus colegas de curso enquanto o realizo.			
No curso, leio todas as respostas a postagens que faço.	0,520		
Primeiro compreendo a estrutura geral do conteúdo do curso para depois estudá-lo em detalhes.			

Ao estudar os conteúdos do curso, busco conhecimentos fundamentados em fatos.	0,546	
Ao estudar os materiais do curso, foco em fatos e dados.	0,699	
Uso minha intuição (entender as coisas no momento, sem necessidade de realizar raciocínios complexos) para aprender o conteúdo do curso.	0,663	
Reflico sobre o conteúdo do curso para melhor compreendê-lo.		
Reflico sobre as possibilidades de aplicação no meu contexto profissional do que aprendi no curso.	0,675	
Ao estudar o material do curso, foco em ideias para aprender.	0,471	

Foram adotados como critérios que nenhum item tivesse carga fatorial inferior a 0,4 ou apresentasse cargas fatoriais superiores a 0,3 em mais de um fator. Na estrutura escolhida, sete itens apresentaram cargas inferiores a 0,4 e dois itens apresentaram cargas fatoriais compartilhadas significativas. Esses itens foram retirados e, em seguida, a simulação de Monte Carlo foi realizada novamente e apontou que a estrutura ideal seria de dois fatores. Além disso, um dos fatores, após a retirada contava com apenas dois itens. Dessa forma, foi realizada uma análise com dois fatores sem os itens que já haviam sido retirados. Nessa segunda análise, foram retirados mais 4 itens usando os mesmos critérios. A estrutura final encontrada contou com 18 itens organizados em 2 fatores. Essa nova estrutura mostrou-se fatorável, com KMO igual a 0,896 (Tabela 7).

Tabela 7

Testes KMO e Bartlett da análise fatorial exploratória após retirada de itens

Medida Kaiser-Meyer-Olkin de adequação de amostragem.		0,896
	Aprox. Qui-quadrado	2040,775
Teste de esfericidade de Bartlett	gl	153
	Sig.	0,000

A estrutura encontrada é a apresentada na Tabela 8. O Fator 1 foi denominado de Abordagem Profunda e o Fator 2 foi denominado de Organização e Potencialização da Aprendizagem.

Tabela 8
Matriz de padrão da análise fatorial exploratória após a retirada de itens

	Fator	
	1	2
Realizo um curso de cada vez.	0,447	
Verifico minhas notas do curso assim que elas são lançadas.	0,429	
Acesso os materiais do curso várias vezes.		0,851
Busco materiais além dos que são passados no curso.		0,897
Divido os materiais de curso de forma a criar um ritmo de estudo que siga até o final do curso.		0,827
Realizo todos os exercícios do curso para praticar o que aprendi.		0,683
Realizo apenas as atividades do curso que são pontuadas.		0,628
Aprendo melhor sozinho os conteúdos do curso.		0,488
Interajo com meus colegas de curso e com o tutor sobre assuntos que não estão diretamente ligados ao conteúdo.	0,672	
Posto minha dúvida quando tenho uma relacionada ao conteúdo do curso.	0,53	
Quando tenho uma dúvida, pergunto a meus colegas de curso.	0,55	
Interajo com o tutor durante o curso.	0,625	
No curso, leio todas as respostas a postagens que faço.	0,483	
Ao estudar os conteúdos do curso, busco conhecimentos fundamentados em fatos.	0,542	
Ao estudar os materiais do curso, foco em fatos e dados.	0,674	
Uso minha intuição (entender as coisas no momento, sem necessidade de realizar raciocínios complexos) para aprender o conteúdo do curso.	0,673	
Reflito sobre as possibilidades de aplicação no meu contexto profissional do que aprendi no curso.	0,658	
Ao estudar o material do curso, foco em ideias para aprender.	0,487	

Os dois fatores apresentaram bons indicadores de consistência interna, sendo os alfas de Cronbach respectivamente 0,85 e 0,87. A correlação entre os fatores é de 0,41 (Tabela 9) e a variância explicada conjuntamente é de 41,3% (Tabela 10).

Tabela 9
Correlação entre os fatores da análise fatorial exploratória após a retirada dos itens

Fator	1	2
1	1,000	0,410
2	0,410	1,000

Tabela 10
Variância explicada e autovalores da análise fatorial exploratória após a retirada de itens

Fator	Somadas de extração de carregamentos ao quadrado			Somas de rotação de carregamentos ao quadrado
	Total	% de variância	% cumulativa	Total
1	5,272	29,289	29,289	4,573
2	2,169	12,052	41,341	4,125

2.3.2 Análise fatorial confirmatória

Uma vez definida a estrutura do questionário a partir dos indícios de validade estatística fatorial, essa estrutura foi testada por meio de análise fatorial confirmatória. Para a amostra aleatória de 310, escolhida da população de respondentes depois de retirados os casos usados na análise exploratória, foram inicialmente analisados os *outliers* univariados. A quantidade somada de casos extremos identificados para cada uma das variáveis foi superior a 20% da amostra, dessa forma, optou-se por não retirar nenhum caso extremo univariado. Os casos extremos multivariados foram identificados pela distância de Mahalanobis. Os 21 casos identificados por esse método foram retirados da amostra, ficando sua versão final com 289 respondentes. A análise de evidências de normalidade apontou que os dados eram assimetricamente distribuídos. Foram testadas transformações para buscar alcançar a normalidade dos dados (p. ex.: raiz quadrada, logaritmo, razão, entre outros). Outra vez, nenhuma delas alterou de forma significativa a distribuição dos dados e, por isso, eles foram analisados sem transformação.

A seguir, o ajuste do modelo foi testado usando como critérios os índices: razão entre qui-quadrado (χ^2) e graus de liberdade (g.l.), Índice de Ajuste Normalizado (NFI), Índice de Tucker Lewis (TLI), Índice Comparativo de Ajuste (CFI), Índice de Qualidade do Ajuste (GFI), Índice Ajustado de Qualidade do Ajuste (AGFI) e Raiz da Média dos Quadrados dos Erros de Aproximação (RMSEA). Os dados encontrados apontam para um bom ajuste do modelo [χ^2 (18, N = 289) = 187,79, $p < 0,001$; $\chi^2/\text{gl} = 1,53$; NFI=0,91; TLI = 0,96; CFI=0,96;

GFI=0,93; AGFI = 0,91; RMSEA = 0,04]. Todos os índices de ajuste do modelo estão adequados. Para chegar a este modelo, foram apontadas onze correlações entre erros, sendo todas elas entre itens de mesmo fator (Figura 8). As cargas fatoriais do fator de Organização e Potencialização da Aprendizagem foram todas iguais ou superiores a 0,6. Para o fator de Abordagem Profunda, houve apenas um item com carga fatorial abaixo de 0,4 (0,39), que foi mantido. Com base nos dados apresentados, entende-se que há evidências confirmatórias de validade estatística da estrutura anteriormente encontrada na análise fatorial exploratória.

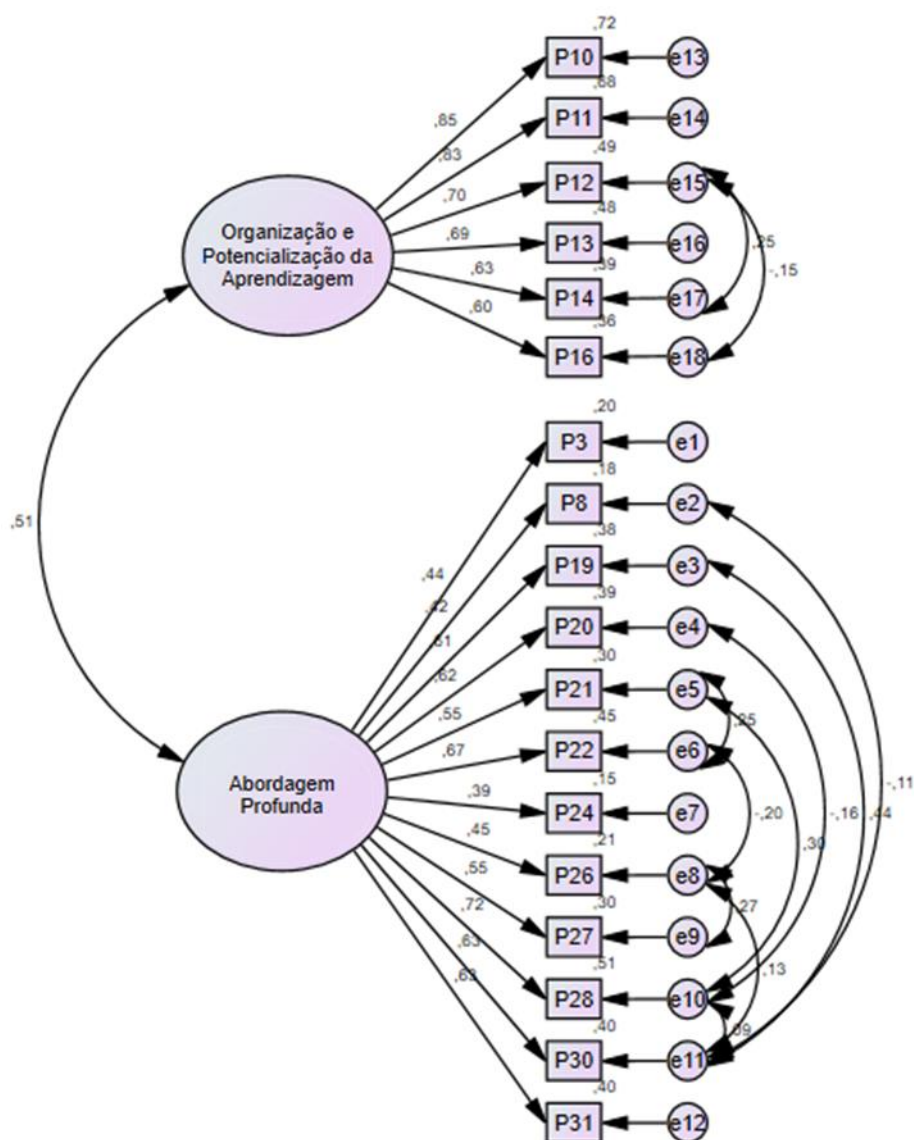


Figura 8. Modelo final da EEA-Ead_C

2.4 Discussão

A escala construída apresentou boas evidências de validade e pode ser usada para o passo seguinte, que é a detecção de estilos de aprendizagem a partir de comportamentos no AVA. Há indícios robustos, advindos tanto da análise exploratória quanto da confirmatória, que os estilos de aprendizagem para quem fez cursos no AVA da instituição em questão agrupam-se em dois fatores. Há diferenças relevantes na emersão dos fatores, quando esta escala, na qual o respondente informa frequência de comportamentos, é comparada às escalas com respostas relativas a preferências de aprendizagem, usadas nos instrumentos de Jesus (2014) e Salles (2007).

O Fator 1, Abordagem Profunda, agregou itens que foram classificados, pelos juízes, como Regulação, Interação Ativa e Passiva e Processamento Cognitivo. Nesse Fator, estão cinco itens de Interação e os outros três foram eliminados. Dos eliminados, dois deles são de Interação Passiva e apenas um item de Interação Passiva foi mantido (Quadro 2). Os itens eliminados referem-se à leitura da interação de terceiros e o item mantido refere-se à leitura de uma postagem que o próprio estudante fez. Sendo assim, foram eliminados itens que se referem à interação com materiais cuja produção foi iniciada por terceiros e mantido o item que se refere à interação do indivíduo com materiais cuja produção foi iniciada por ele próprio. Entende-se que a divisão teórica feita para a interação não é a mais adequada. O Fator 1 contém apenas o item que descreve o estudante como ativo em seu processo de aprendizagem, isto é, que faz postagens e lê reações a elas, o que parece estar alinhado com a divisão proposta por Berings, Poell e Simons (2005).

Quadro 2

Itens mantidos e retirados do fator teórico de Interação Passiva

Fator 1 – Itens de Interação Passiva	
<i>Item mantido</i>	
No curso, leio todas as respostas a postagens que faço.	
<i>Itens eliminados</i>	
Leio todas as perguntas, respostas e comentários que meus colegas e tutor postam no curso.	
Leio todos os tópicos que são postados nos fóruns do curso.	

Os itens de Processamento Cognitivo do Fator 1 foram quase todos mantidos, com a exceção de dois (Quadro 3). Dos itens retirados, um deles é similar a um item do Fator 1 da escala de Jesus (2014): “Primeiro compreendo a estrutura geral do conteúdo do curso para depois estudá-lo em detalhes.” Essa é uma forma cognitiva de organizar o estudo e difere dos itens mantidos, pois eles parecem ser estratégias cognitivas para atingir a aprendizagem ou a aplicação dela. Essas estratégias cognitivas podem representar a busca ativa pela aprendizagem e, se analisadas em termo de preferência, podem ser excludentes. Por exemplo, um estudante pode preferir focar em ideias ou em fatos e dados para aprender. Uma vez que as cargas fatoriais são todas positivas, essas são estratégias que os estudantes usam mais ou menos e não preferências opostas. O uso de várias estratégias cognitivas para aprender um conteúdo está ligado a uma abordagem profunda, enquanto que o uso de poucas estratégias está ligado a uma abordagem superficial (Entwistle & Peterson, 2004).

Quadro 3

Itens mantidos e retirados do fator teórico de Processamento Cognitivo

Fator 1 - Itens de Processamento Cognitivo
<i>Itens mantidos</i>
Ao estudar os conteúdos do curso, busco conhecimentos fundamentados em fatos. Ao estudar os materiais do curso, foco em fatos e dados. Uso minha intuição (entender as coisas no momento, sem necessidade de realizar raciocínios complexos) para aprender o conteúdo do curso. Reflito sobre as possibilidades de aplicação no meu contexto profissional do que aprendi no curso. Ao estudar o material do curso, foco em ideias para aprender.
<i>Itens eliminados</i>
Primeiro compreendo a estrutura geral do conteúdo do curso para depois estudá-lo em detalhes. Reflito sobre o conteúdo do curso para melhor compreendê-lo.

A divisão teórica proposta para a categoria de Regulação da aprendizagem, feita pelos juízes, não foi confirmada. Os dois itens dessa categoria 1 podem ser entendidos como um desdobramento de uma abordagem profunda (Quadro 4). O estudante interessado em realmente aprender o conteúdo normalmente realiza apenas um curso de cada vez e tem interesse na nota que recebe em uma avaliação. Da mesma forma, tem interesse específico em conhecer reações a suas postagens, razão pela qual aquele único item de interação, anteriormente discutido aqui, foi o único que permaneceu no Fator 1. Portanto, tal item provavelmente é mais um caso de regulação e não a interação originalmente proposta.

Quadro 4

Itens mantidos e retirados do fator teórico de Regulação

Itens de Regulação	
<i>Itens mantidos</i>	<i>Fator</i>
Realizo um curso de cada vez.	1
Verifico minhas notas do curso assim que elas são lançadas.	1
Acesso os materiais do curso várias vezes.	2
Busco materiais além dos que são passados no curso.	2
Divido os materiais de curso de forma a criar um ritmo de estudo que sigo até o final do curso.	2
Realizo todos os exercícios do curso para praticar o que aprendi.	2
Realizo apenas as atividades do curso que são pontuadas.	2
<i>Itens eliminados</i>	
Logo antes das avaliações do curso, reviso os materiais de estudo.	
Realizo todas as atividades do curso, mesmo as que não são pontuadas.	
Acesso os materiais adicionais do curso.	
Defino minhas próprias metas para estudar os materiais do curso.	
Entrego com antecedência todas as tarefas do curso que têm prazo de entrega.	
Leio todos os feedbacks que o tutor passa sobre as tarefas que realizei no curso.	
Acesso os materiais do curso na ordem que são apresentados.	

Considerando o conjunto dos itens que permaneceram no Fator 1, entende-se que o Fator 1 (Quadro 5) é um fator que expressa a abordagem profunda (Entwistle, 1981). Ele separa os estudantes entre os que se esforçam para realmente entenderem o conteúdo (abordagem profunda) e os que prioritariamente desejam ser aprovados (abordagem superficial). Esses conceitos são similares, respectivamente, aos conceitos de direcionamento para o sentido e direcionamento para a reprodução de Vermunt (Entwistle & Peterson, 2004).

Quadro 5

Itens finais do Fator 1

Fator 1 - Abordagem Profunda
<ul style="list-style-type: none"> - Ao estudar os materiais do curso, foco em fatos e dados. - Uso minha intuição (entender as coisas no momento, sem necessidade de realizar raciocínios complexos) para aprender o conteúdo do curso. - Interajo com meus colegas de curso e com o tutor sobre assuntos que não estão diretamente ligados ao conteúdo. - Reflito sobre as possibilidades de aplicação no meu contexto profissional do que aprendi no curso. - Interajo com o tutor durante o curso. - Quando tenho uma dúvida, pergunto a meus colegas de curso. - Ao estudar os conteúdos do curso, busco conhecimentos fundamentados em fatos. - Posto minha dúvida quando tenho uma relacionada ao conteúdo do curso. - Ao estudar o material do curso, foco em ideias para aprender. - No curso, leio todas as respostas a postagens que faço. - Realizo um curso de cada vez. - Verifico minhas notas do curso assim que elas são lançadas.

O Fator 2, Organização e Potencialização da Aprendizagem (Quadro 6), agrupa cinco itens que eram hipotetizados como de regulação e o item “Aprendo melhor sozinho.” Ele agrupa itens das categorias Potencialização da Efetividade e Estudo Organizado, que podem ser classificados como fatores da abordagem estratégica. Na abordagem estratégica, também há as categorias de gerenciamento do tempo e motivação para o resultado (Entwistle *et al.*, 2013), que não foram encontrados na EEA-Ead_C. A abordagem estratégica envolve vários comportamentos autorregulatórios, que emergiram em um fator, assim como em Jesus (2014). Provavelmente as características dos cursos disponibilizados no AVA influenciaram no agrupamento e exclusão dos comportamentos regulatórios, uma vez que eles são muito dependentes do contexto.

Quadro 6

Itens finais do Fator 2

Fator 2 - Organização e Potencialização da Aprendizagem
Busco materiais além dos que são passados no curso.
Acesso os materiais do curso várias vezes.
Divido os materiais de curso de forma a criar um ritmo de estudo que sigo até o final do curso.
Realizo todos os exercícios do curso para praticar o que aprendi.
Realizo apenas as atividades do curso que são pontuadas.
Aprendo melhor sozinho os conteúdos do curso.

A princípio, o questionário de estilo de aprendizagem proposto nessa dissertação não é generalizável. É necessária a aplicação de instrumentos similares em outros contextos para investigar como o contexto afeta o agrupamento dos comportamentos analisados. A estrutura tem bons indícios de validade estatística, no entanto, indica-se que, ao usar a mesma estrutura teórica, os itens sejam revistos para que sejam adequados ao ambiente virtual a que se referem.

CAPÍTULO 3 – ESTILOS DE APRENDIZAGEM AUTODECLARADOS E OBSERVADOS

O Capítulo 3 abordará a relação entre os estilos de aprendizagem autodeclarados e observados. A escala desenvolvida no Capítulo 2 pautou a definição usada para estilos de aprendizagem e conseqüentemente os comportamentos que são entendidos como expressões daqueles estilos. Os comportamentos serão transformados em indicadores com base em *footprints* de um AVA e esses indicadores serão usados para identificar os estilos autodeclarados (objetivo específico 3). Antes de usar os indicadores, é necessário analisar se eles são viáveis para essa identificação, que significa analisar sua constância em diferentes contextos (objetivo específico 2). Além disso, eles serão agrupados, de forma a criar os estilos de aprendizagem observados. Para realizar esses procedimentos serão analisadas as áreas de *Educational Data Mining* e *Learning Analytics* e será revisada a literatura de detecção automática de estilos de aprendizagem. Em seguida serão expostos o método e os resultados para a classificação dos estudantes de acordo com os estilos autodeclarados e o teste do instrumento de detecção automática. Por último, os estilos de aprendizagem autodeclarados e observados serão comparados (objetivo específico 4) e os resultados serão discutidos.

Diante dos apontamentos dos estudos de ferramentas automáticas de detecção de estilos de aprendizagem, foram revisados os instrumentos disponíveis tanto de forma teórica quanto prática partindo do ponto de vista da psicologia. Essa revisão foi feita para escolher o conceito de estilo de aprendizagem que seria usado e a ferramenta. Foi com base nessa revisão que se identificou a necessidade de criação de uma nova ferramenta baseada em uma teoria que não foi usada por outros autores.

3.1 Educational Data Mining e Learning Analytics

As duas áreas que surgiram para analisar *footprints* foram o *educational data mining* (EDM) e o *learning analytics* (LA). O EDM pode ser entendido como a aplicação de técnicas de *data mining*, ou mineração de dados, em dados específicos que vêm de ambientes educacionais para a solução de problemas educacionais (Romero & Ventura, 2010). Essa área foi criada há poucos anos, mas a produção e a quantidade de revistas e *workshops* específicos deste tema têm crescido exponencialmente (Romero & Ventura, 2013). Os dados analisados não estão restritos ao contato do estudante com o ambiente, pois também incluem a colaboração entre eles em *chats* e fóruns, dados administrativos de professor, tutor, escola onde estuda, setor onde trabalha e dados demográficos. Esses dados têm uma organização hierárquica (e.g., módulo do curso, assunto, tarefa), estão dentro de um determinado contexto (e.g., é atribuída uma tarefa a um dado estudante em uma dada turma), partem do nível *micro* (são gravados no nível do estudante e podem ser agregados a partir do nível mais baixo) e são longitudinais (os dados são gravados ao longo do tempo e há a informação de quando foram gravados). Essa área surgiu principalmente da junção entre ciência da computação, educação e estatística, como representa a Figura 9.

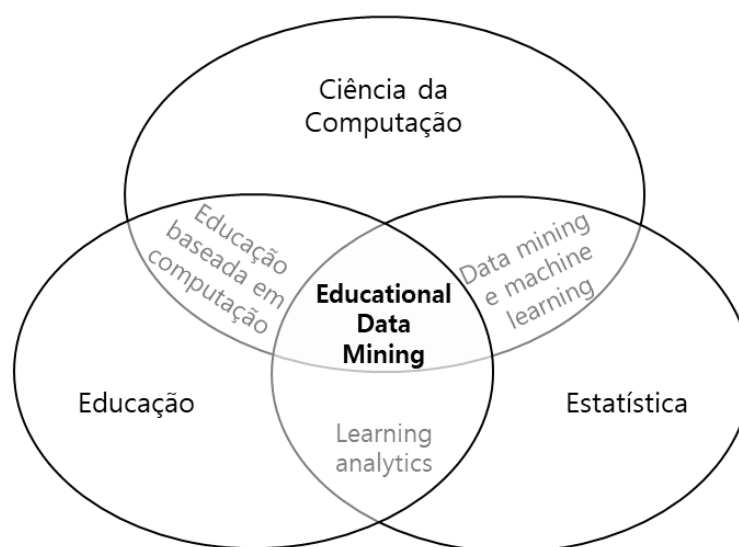


Figura 9. Áreas relacionadas a *Educational Data Mining*.

Fonte: Romero, C., & Ventura, S. (2013). *Data mining in education*. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12-27.

O LA é outra área que surgiu nos últimos anos e que é muito relacionada a EDM. Refere-se à mensuração, coleta, análise e relato de dados sobre aprendizes em seus contextos, com o objetivo de entender e otimizar o aprendizado e o ambiente em que ele ocorre. A mencionada área está na intersecção entre os campos técnico e social das teorias de aprendizagem. As ferramentas desenvolvidas demandam grande *expertise* técnica, mas seus impactos são sentidos no sistema social de aprendizagem. Por esse motivo, o foco dos profissionais de LA é criar uma ponte entre as duas comunidades (Siemens & Gasevic, 2012).

Apesar de próximas, EDM e LA partem de vertentes diferentes e têm 5 diferenças principais: (i) preferência por paradigmas automatizados ao levantar informações (EDM) em relação ao uso de técnicas para auxiliar no levantamento de informações pelos profissionais (LA); (ii) foco em reduzir os fenômenos em seus componentes e analisar cada um deles (EDM) em relação a foco em entender o sistema como um todo, considerando toda sua complexidade (LA); (iii) foco na adaptação automatizada dos recursos (EDM) em relação a fornecer aos profissionais as informações necessárias para realizarem as adaptações (LA); (iv) origem em *softwares* educacionais e modelagem de comportamento de estudantes, com foco em prever resultados do curso (EDM) em relação a origem em *web* semântica e intervenção sistêmica; e (v) usa métodos como clusterização e modelagem Bayesiana (EDM) em relação a métodos como *social network analysis* e *sentiment analysis* (Siemens & Baker, 2012).

A área de EDM é muito nova e 1995 é apontado como o ano de seu início. Em 2000 surgiu o primeiro *workshop* da área e desde então houve a produção de vários livros, criação de periódicos específicos para a área e comunidades de pesquisadores. A área de LA, que também é conhecida como *learning analytics and knowledge* (LAK), é mais recente, sendo que a primeira conferência da área ocorreu em 2011 (Figura 10). A projeção é que até o ano de 2022 todas as pesquisas da área de educação envolvam *analytics* e mineração de dados (Baker & Inventado, 2014).

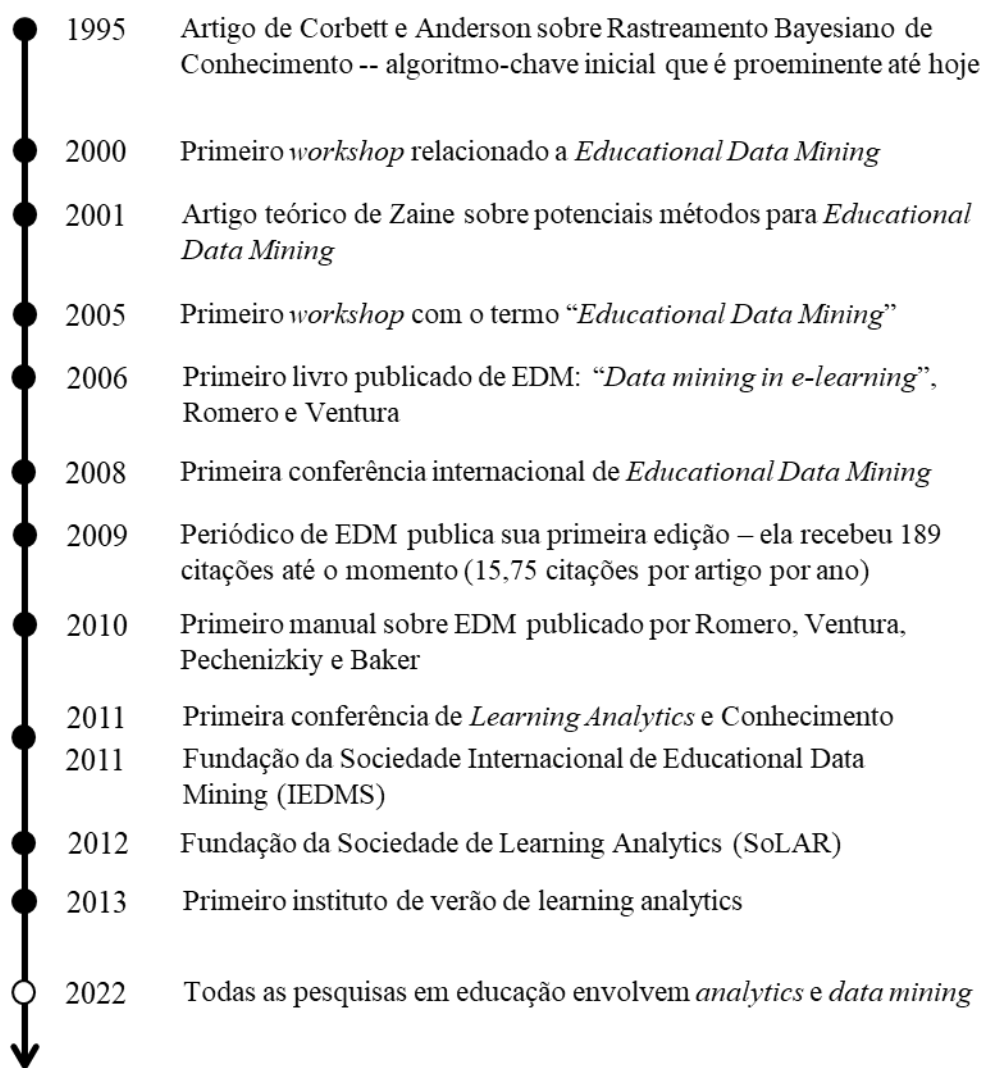


Figura 10. Linha do tempo de acontecimentos relevantes de EDM.

Fonte: Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In Learning analytics (pp. 61-75). Springer, New York, NY.

As principais aplicações de EDM e LA, em ensino a distância (EAD), são: (i) predição de desempenho de estudantes; (ii) detecção de risco de evasão desses estudantes; (iii) visualização de dados; (iv) *feedback* inteligente para estudantes, tutores e planejadores; (v) recomendação de cursos; (vi) estimativa de habilidade de estudantes; (vi) detecção de comportamentos de estudante; (vii) colaboração entre estudantes e em grupos; (viii) análise de suas redes sociais; (ix) desenvolvimento de mapas de conceito; (x) construção de trilhas de aprendizagem; (xi) planejamento e agendamento. Destes tópicos, o que gerou maior interesse

foi a detecção de comportamentos, que é criar um modelo a partir dos comportamentos do estudante no ambiente virtual (Sin & Muthu, 2015).

As principais forças dessas duas novas áreas são o grande volume de dados disponíveis em diversas organizações e ambientes; o uso de métodos algorítmicos que são válidos e já existem; as múltiplas formas de visualização dos estudantes e professores que são disponibilizadas, os modelos de estudantes cada vez mais precisos para guiar a adaptação e a personalização dos sistemas; a identificação de momentos e padrões críticos relacionados à aprendizagem; e *insights* sobre estratégias de aprendizagem e comportamentos. As principais oportunidades são a possibilidade de tornar compatíveis as diferentes ferramentas e aplicações na forma de uma plataforma; oportunidades de aprendizagem multimodais e afetivas baseadas em métricas sofisticadas; autorreflexão, autoconsciência, autoaprendizagem em sistemas massivos que são autônomos e inteligentes; resultados advindos de EDM e LA em formatos compatíveis com outros sistemas para que possam ser usados como suporte à tomada de decisão. As fraquezas são a interpretação incorreta dos resultados; as bases de dados heterogêneas e não compatíveis; a ausência de bons resultados qualitativos; excesso de informação; incerteza sobre a relevância dos sistemas para os profissionais da área, pois atualmente só os que possuem muita habilidade em análise de dados conseguiram interpretá-los. As ameaças são os problemas éticos, principalmente com a privacidade dos dados; a falta de generalidade dos resultados; a possibilidade de classificação incorreta dos padrões; e achados contraditórios durante a implementação (Papamitsiou & Economides, 2014) (Quadro 7).

Quadro 7

Matriz SWOT das pesquisas em Learning Analytics e Educational Data Mining

Forças	Fraquezas
<ul style="list-style-type: none"> Grande volume disponível de dados educacionais -> acurácia crescente de resultados experimentais. Uso de métodos algorítmicos preexistentes, poderosos e válidos. Múltiplas formas de visualização para apoiar estudantes e professores. Modelos do usuário mais precisos para guiar a adaptação e personalização dos sistemas. Revela momentos críticos e padrões de aprendizagem. Possibilidade de insights sobre estratégias de aprendizagem e comportamento. 	<ul style="list-style-type: none"> Interpretação equivocada dos resultados devido a fatores do julgamento humano - foco em reportar e não em decidir. Fontes de dado heterogêneas: não há um vocabulário descritivo unificado - problemas com representação dos dados. Resultados de pesquisa majoritariamente quantitativos. Métodos qualitativos ainda não obtiveram resultados significativos. Excesso de informação - sistemas complexos. Incerteza: "já estamos prontos?" Até o momento, apenas
Oportunidades	Ameaças
<ul style="list-style-type: none"> Uso de <i>Open Linked Data</i> para padronização e compatibilização dos dados entre diferentes ferramentas e aplicações -> desenvolvimento de plataforma generalizada. Oportunidades de aprendizagem multimodais e afetivas baseadas em métricas sofisticadas. Autorreflexão / Autoconhecimento / Autoaprendizagem em sistemas massivos inteligentes e autônomos. Fornecer resultados legíveis por máquina de LA e EDM para outros sistemas de tomada de decisão baseados em dados. Modelo de aceitação: p.ex., utilidade percebida, expectativa de objetivos, jogabilidade percebida, confiança, etc. 	<ul style="list-style-type: none"> Dilemas éticos - privacidade dos dados. Over-analysis: profundidade excessiva das análises com perda de generabilidade. A abordagem de excesso de granularidade pode ameaçar a visão holística; olhar a árvore e não perceber a floresta. Possibilidade de classificação incorreta de padrões. Confiança: achados contraditórios durante a implementação.

Fonte: Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society*, 17(4).

Os comportamentos em AVA foram citados por Abbad, Zerbini e Souza (2010), que apontaram a grande quantidade de dados armazenados no Brasil. No entanto, não foi encontrado qualquer estudo no Brasil que usou as informações armazenadas de comportamentos do indivíduo em AVA de forma sistemática e inferindo construtos psicológicos a partir desses comportamentos. Unindo isso aos estilos de aprendizagem, há uma linha de pesquisa relativamente nova, que é a detecção dos estilos de aprendizagem a partir de *footprints* de AVA. Nas seções seguintes, a literatura da área foi analisada usando como base revisões de literatura internacionais e elas foram complementadas por uma revisão própria.

3.2 Detecção Automática de Estilos de Aprendizagem

Uma das variáveis do estudante analisada em AVAs é o construto estilo de aprendizagem. Existem muitas definições de estilos de aprendizagem e não há consenso na literatura (Coffield *et al.*, 2004). Os estilos podem ser entendidos como:

“um composto de características cognitivas, afetivas e fatores psicológicos que servem como um indicador relativamente estável de como um aprendiz percebe, interage com e responde ao ambiente de ensino.” (Keefe, 1979 em Feldman, Monteserin & Amandi, 2015, p. 1).

Existiriam 71 modelos de estilos de aprendizagem considerados relevantes, que são usados para classificar os indivíduos de acordo com a teoria, segundo Coffield *et al.* (2004). Os estilos variam tanto em termos teóricos quanto metodológico e há sobreposição entre esses modelos. De toda forma, estilo cognitivo é um construto válido que ajuda a entender o processo de aprendizagem dos indivíduos e vários desses modelos têm relação com a aprendizagem em AVAs (Feldman, Monteserin & Amandi, 2015).

A forma tradicional de identificar os estilos de aprendizagem é por meio de escalas, conforme foi apresentado no Capítulo 2, o que apresenta vários problemas, conforme discutido no Capítulo 1. Além dos problemas que são aplicáveis a qualquer investigação de construto psicológico por meio de escalas, os estilos de aprendizagem apresentam mais dois problemas específicos. O estudante muitas vezes não conhece suas preferências de aprendizagem e o estilo de aprendizagem pode variar com o tempo e com o contexto. Por isso, uma solução buscada por vários autores (Al-Azawei & Badii, 2014; Feldman *et al.*, 2015; Oliveira, Sales, Pereira & Moreira, 2018; Jaiswal, Singh & Ahuja, 2017) foi identificar o estilo de aprendizagem de forma automática analisando os *footprints* em AVAs. As vantagens dessa abordagem são que os estilos são inferidos a partir do próprio comportamento do aprendiz, não é necessário qualquer trabalho adicional, como o preenchimento de questionários, e as informações são colhidas ao longo de toda interação do aprendiz com o ambiente e não apenas em um período específico. A identificação automática permite que a informação seja colhida sem qualquer intervenção (Feldman *et al.*, 2015).

Conhecimento sobre um determinado estudante é chamado de modelo do estudante (ME). A criação de um ME pode envolver todas as variáveis dele disponíveis no AVA (Brusilovsky, 1994). O processo de transformar as informações brutas do sistema em um ME é chamado de modelagem do estudante (Self, 1994). O ME pode considerar muitas variáveis, como, por exemplo, conhecimento anterior, experiência, interesses, objetivos, estilos de aprendizagem, traços cognitivos e estratégias de aprendizagem. O foco das pesquisas dessa área tem sido em estilos de aprendizagem principalmente porque há evidências empíricas de que o processo de aprendizagem pode ser melhorado se for adaptado ao estilo de aprendizagem do estudante e porque é uma variável relativamente estável no tempo (Al-Azawei & Badii, 2014).

Com isso, as pesquisas em detecção de estilos de aprendizagem têm crescido muito desde a primeira década do presente século (Al-Azawei & Badii, 2014; Feldman *et al.*, 2015; Oliveira, Sales, Pereira & Moreira, 2018; Jaiswal, Singh & Ahuja, 2017). As pesquisas sobre esse tema são focadas na detecção do estilo para que sejam usados em Sistemas Educacionais Adaptativos em Hipermedia (SEAH). Os SEAHs apresentam o conteúdo e moldam o contexto do curso de acordo com variáveis do estudante. Foram encontradas 4 revisões de literatura sobre a detecção dos estilos de aprendizagem, sendo que duas delas dirigiram seu foco para a detecção dos estilos de aprendizagem (Feldman *et al.*, 2015; e Oliveira, Sales, Pereira & Moreira, 2018) e as outras duas tenderam a analisar a detecção dos estilos com o foco em SEAHs (Al-Azawei & Badii, 2014; e Jaiswal, Singh & Ahuja, 2017). Elas serão abordadas na próxima seção.

3.2.1 Revisão de Literatura

Para essa Dissertação, além da análise das revisões, foi realizada uma revisão das publicações em detecção de estilos de aprendizagem entre 2015 e 2018 para complementar a revisão de Feldman *et al.* (2015). Em todas as publicações analisadas, a detecção de estilos de

aprendizagem foi usada como um fim em si ou para o uso em SEAHs. Não foi encontrado qualquer outro fim para o qual a detecção automática tenha sido utilizada. Isto é, foram confirmadas as duas tendências já identificadas nas quatro revisões anteriormente citadas.

A detecção de preferências de estilos de aprendizagem pode ser dividida em detecção explícita e implícita. Na explícita, o estilo é determinado perguntando diretamente para o indivíduo e o instrumento usado para isso são as escalas. Essa abordagem também pode ser chamada de modelagem guiada, *feedback* explícito do usuário e abordagem colaborativa (Al-Azawei & Badii, 2014). A implícita é a detecção automática e pode ser baseada em dados (DBD) ou em teoria (DBT). A detecção baseada em dados usa *footprints* dos estudantes para identificar os estilos a partir de comportamentos em AVA. Constrói um modelo que busca imitar as respostas ao questionário. Sua principal vantagem é que pode ser muito precisa, por usar comportamentos observáveis (ou registrados nos sistemas), mas tem como desvantagens depender exclusivamente dos dados disponíveis, que podem ser insuficientes, e de depender da validade do questionário cujas respostas foram imitadas. A DBT determina os comportamentos que são preferidos pelos estudantes a partir da literatura científica. A DBT usa os comportamentos como pistas para o estilo de aprendizagem e cria regras para ligar os comportamentos aos estilos. Os comportamentos de cada estudante são analisados de acordo com as regras e isso determina o estilo de aprendizagem (Graf, 2007). Na Figura 11 são apresentadas as relações entre os modelos de estilos de aprendizagem e as abordagens para sua detecção automática. No Quadro 8 são apresentados os prós e contras de cada uma das abordagens.

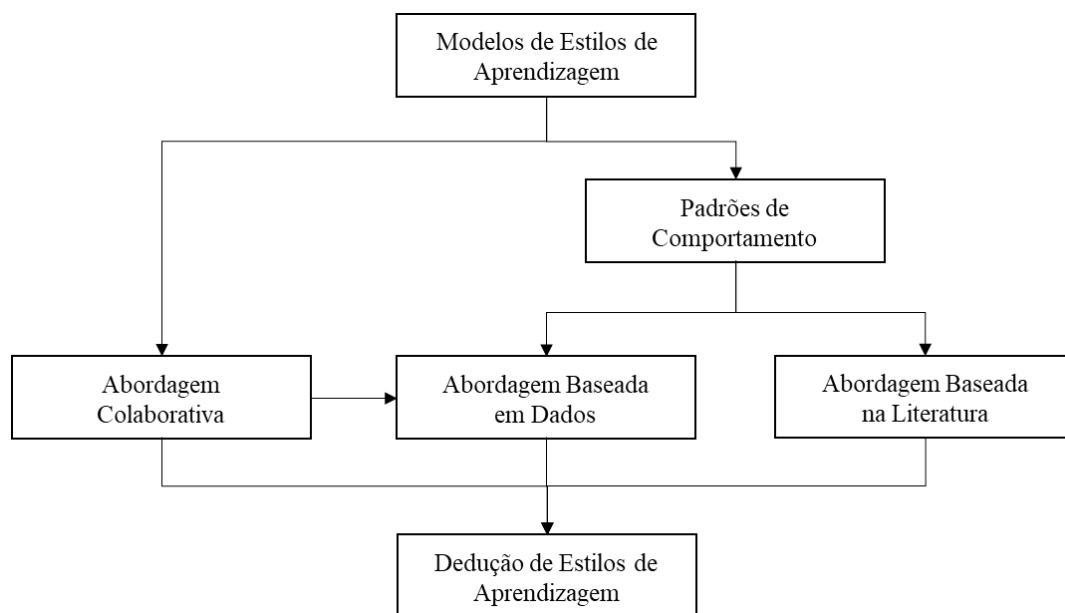


Figura 11. Relação entre abordagens e modelos de estilos de aprendizagem.

Fonte: Al-Azawei, A., & Badii, A. (2014). State of the art of learning styles-based adaptive educational hypermedia systems (LS-BAEHSs). *International Journal of Computer Science & Information Technology*, 6(3), 1-19.

Quadro 8

Prós e Contras das abordagens colaborativa, baseada em dados e baseada na literatura

	Abordagem colaborativa	Abordagem baseada em dados	Abordagem baseada na literatura
Prós	Fornecer dados coletados como auto expressão autêntica.	Reflete os comportamentos naturais dos estudantes.	Reflete os comportamentos naturais dos estudantes.
	Menos ruídos e dados espúrios.	Mais preciso em representar suas preferências.	Mais preciso em representar suas preferências.
	Dado pode ser extraído em formato estruturado e padronizado.	É um processo dinâmico, o que significa que pode ser usado para criar modelos do estudante do zero e atualizá-los.	É um processo dinâmico, o que significa que pode ser usado para criar modelos do estudante do zero e atualizá-los.
			Depende unicamente dos comportamentos e ações do estudante no AVA.
Contras	Usuários podem não conseguir expressar suas preferências diretamente.	Alta complexidade e custo computacional.	Alta complexidade e custo computacional.
	Respostas arbitrárias podem ser dadas caso as perguntas sejam confusas ou o questionário longo e as respostas estão sujeitas a vieses.	Dificuldade de medir e interpretar os comportamentos dos estudantes.	Dificuldade de medir e interpretar os comportamentos dos estudantes.
	Dados são estáticos, enquanto que preferências podem mudar.	O processo de classificação de aprendizagem e estilo cognitivo é <i>offline</i> .	O processo de classificação de aprendizagem e estilo cognitivo é <i>offline</i> .
	Pode ser percebida como perturbadora, problemática e demorada pelo usuário.	A acurácia dos resultados depende unicamente dos dados disponíveis e da identificação dos padrões de comportamentos.	

Fonte: Al-Azawei, A., & Badii, A. (2014). State of the art of learning styles-based adaptive educational hypermedia systems (LS-BAEHSs). *International Journal of Computer Science & Information Technology*, 6(3), 1-19.

Outro ponto a ser analisado sobre a detecção dos estilos é se ela é estática ou dinâmica. A estática é quando a detecção é realizada apenas uma vez, não sendo atualizada com o tempo. Na dinâmica, os estilos são detectados mais de uma vez, alterando de acordo com a mudança de comportamento do estudante. Há um consenso de que os estilos são relativamente fixos ao longo do tempo, mas que podem variar. Portanto, a forma mais precisa de aferi-los seria a dinâmica, no entanto, este método apresenta muitos desafios metodológicos e tecnológicos. Ressalta-se que os modelos dinâmicos não reavaliam

constantemente os estilos, mas que registram os comportamentos de forma longitudinal e que periodicamente revisam os estilos (Dorça, 2012).

As revisões de literatura de detecção de estilos de aprendizagem analisam principalmente o modelo de estilo de aprendizagem, o ME, as técnicas de análise de dados usadas, o AVA e a eficácia da ferramenta. As pesquisas que criaram formas de detecção automática de estilos de aprendizagem foram objeto de análise de Feldman *et al.* (2015). Em 27 artigos publicados de 1999 a 2012, as pesquisas usaram 4 modelos de diferentes de estilos de aprendizagem, o de Biggs (Biggs, 1987), o de Felder-Silverman (Felder & Silverman, 1988), o de Gardner (Gardner, 1993), e o de Kolb (Kolb, 1984), tendo sido o modelo de Felder-Silverman o mais utilizado. Além desses modelos, cinco estudos usaram modelos customizados. As técnicas de análise utilizadas foram redes Bayesianas, árvores de decisão, redes neurais, modelo oculto de Markov, algoritmo genético, raciocínio baseado em casos e modelos probabilísticos gráficos. Delas, a mais utilizada foi a de redes Bayesianas. Quanto ao tipo de informação usada para a construção do ME, elas foram divididas em desempenho, *feedback* e comportamento (Popescu, 2009). O desempenho é a nota que o estudante recebe nas avaliações e foi interpretado de forma que boas performances indicam um estilo adequado ao curso, enquanto que performance ruim indica um estilo não adequado ao curso. O *feedback* é quando o sistema fazer perguntas ao aprendiz e o modelo é ajustado a partir das respostas.

O comportamento são as ações realizadas no AVA. Dentre aqueles estudos analisados, 21 basearam a automatização da identificação dos estilos no comportamento. As variáveis analisadas podem ser divididas em temporais, tentativas e navegação (Stathacopoulou *et al.* 2005). Exemplos de variáveis temporais são o tempo gasto para ler materiais, tempo para encontrar a resposta correta e o tempo total em uma tarefa. As variáveis de tentativa referem-se à quantidade de respostas dadas até que fosse encontrada a correta e a quantidade de vezes que um assunto foi revisto. As variáveis de navegação incluem a quantidade de tópicos

selecionados e o número de vezes que o aprendiz vai para um novo tópico. Alguns estudos usaram variáveis de todos os tipos, mas também há estudos que usaram apenas um tipo delas. As ferramentas de identificação automática de estilos apresentaram boa validade e consistência interna, sendo possível usá-los em sistemas educacionais. Todos os estudos desenvolveram a automatização da identificação do estilo para apenas um curso e a quantidade de participantes variou de 10 a 127. Essas amostras usadas nos estudos são pequenas e a maior parte dos estudos estudou apenas estudantes de graduação de ciência da computação (Feldman et. al., 2015).

Quanto ao tipo de AVA utilizado, todos os estudos criaram um sistema personalizado para a detecção dos estilos. São exemplos destes sistemas: CS383 (Carver, Howard & Lane, 1999), LSAS (Bajraktarevic, Hall & Fullick, 2003), INSPIRE (Papanikolaou, Grigoriadou, Kornilakis & Magoulas, 2003), TANGOW (Paredes & Rodrigues, 2004), AHA! (Stash, Cristea & De Bra, 2006), TSAL (Tseng, Chu, Hwang, & Tsai, 2008), WELSA (Popescu, 2008), com exceção de Graf, Kinshuk e Liu (2008), que usaram o DeLes, que é um *add-on* do *Moodle*. Uma análise mais completa dos sistemas utilizados em detecção automática de estilos de aprendizagem é apresentada por Graf (2007) (Quadro 9). O uso de sistemas personalizados deve-se também à especificidade dos dados do curso e do AVA, mas principalmente porque todos os sistemas citados também são SEAHs. Por isso, também foram analisadas duas revisões de literatura específicas sobre esse tema.

Quadro 9

Sistemas adaptativos de hipermídia em relação aos estilos de aprendizagem (Graf, 2007)

Sistema	Ano	Modelo de Estilo de Aprendizagem	Abordagem para modelagem do estudante	Métodos para prover adaptatividade
CS383	1999	Sensitivo/intuitivo, visual/verbal e sequencial/global, dimensões do FSLM	Questionário <i>Inventory of Learning Styles</i>	Ordenação dos objetos de multimídia
MANIC	2000	Combinação de preferências de aprendizagem	Abordagem automática usando o classificador <i>Naïve Bayes</i> e dados populacionais	Uso de <i>stretchtext</i> (esconder e apresentar conteúdos adicionais)
IDEAL	2002	Determinado pelo professor	Questionário do modelo de estilo de aprendizagem escolhido	Ordenação, inclusão e seleção dos materiais de aprendizagem
MANSPLANG	2002	FSLM	<i>Index of Learning Style</i> para inicializar um processo de racionalização caso a caso para otimização	Escolha de formatos de mídia, estratégias instrucionais e ferramentas de navegação
LSAS	2003	Sequencial/global do FSLM	<i>Index of Learning Style</i>	Esconder/apresentar <i>links</i> e elementos do curso
iWeaver	2003	Preferências de apresentação e psicológicas baseadas no modelo de Dunn e Dunn	Building Excellence Inventory; a abordagem automática é planejada	Ordenação e apresentação de <i>links</i> para selecionar diferentes modos de apresentação e ferramentas de aprendizagem
INSPIRE	2003	Modelo de Honey e Mumford	Questionário de Honey e Mumford para inicializar/atualizar o modelo do estudante manualmente	Método e ordenação da apresentação de conteúdo
TANGOW	2004	Sensitivo/intuitivo e sequencial/global, dimensões do FSLM	<i>Index of Learning Styles</i> para iniciar e abordagem automática da modelagem do estudante revisar as informações no modelo do estudante	Ordem das tarefas e dos elementos nelas
AHA!	2005/ 2006	Determinado pelo professor	Iniciado manualmente e atualizado pelas meta-estratégias instrucionais determinadas	Seleção dos itens apresentados, ordenação da informação e criação de diferentes trilhas de navegação

Fonte: Graf, S. (2007). *Adaptivity in learning management systems focussing on learning styles*. Tese de Doutorado, Faculty of Informatics, Vienna University of Technology, Viena, Áustria.

Uma análise predominantemente descritiva, de 78 artigos publicados de 2001 a 2016, foi feita na revisão de Jaiswal, Singh e Ahuja (2017). A maior parte dos trabalhos é focada nos estudantes (52), no entanto, também há trabalhos focados no professor (7) e mistos (4). Em geral, são realizados em cursos de graduação (53), principalmente de ciência da computação e engenharia (54%). Quanto aos ME, a maior parte é estática (68%), mas os modelos dinâmicos já são bem representativos (32%). A grande maioria (87%) usou questionários como instrumentos para identificação dos estilos. Há preferência pela detecção baseada em dados em relação à detecção baseada em teoria. Quanto aos modelos de estilo de aprendizagem, a maioria (68%) usou o de Felder-Silverman. Esses dados foram resumidos na Tabela 11. A conclusão é que os estilos de aprendizagem detectados automaticamente trazem bons resultados para os SEAHs e que os métodos utilizados são adequados. Dessa forma, indica-se o questionário de Felder-Silverman como uma boa ferramenta, mas que pode ser complementada por outras, uma vez que questionários como o de Kolb trazem outras informações sobre o estudante. Sugere que sejam criados algoritmos que identifiquem modelos mistos de estilos de aprendizagem.

Tabela 11
 Resultados da revisão de literatura em Jaiswal et al. (2017)

	Frequência (f)	%		Frequência (f)	%
1. Tipo dos participantes (n=63)			5. Ferramentas de coleta de dados (n=134)		
Estudante	52	82,54	Questionário de estilo de aprendizagem	55	41,04
Professor / Educador	7	11,11	Teste de progresso da aprendizagem	26	19,40
Misto	4	6,34	Questionário online	23	17,16
2. Nível dos participantes (n=73)			Log em sistema	14	10,45
Ensino Superior / Engenharia	53	72,60	Por meio de entrevista	7	5,22
Ensino Médio	6	8,22	Inventário de estilos cognitivos	7	5,22
Ensino Fundamental	4	5,50	Outro inventário de inteligência	2	1,50
Misto	10	19,69			
3. Modelagem do estudante (n=59)			6. Área / campo do sujeito (n=87)		
Estático	40	67,79	Ciência da computação / engenharia	47	54,02
Dinâmico	19	32,20	Aritmética	11	12,64
4. Ferramentas para modelagem dinâmica (n=34)			Química	7	8,10
Acompanhar o comportamento do estudante	9	26,47	Ciências de gestão	6	6,89
Resultado de testes	7	20,58	Administração	3	3,45
Feedback do estudante	6	17,65	Ciências sociais	2	2,30
Tempo gasto	4	11,76	Direito	2	2,30
Seleção do estudante com base no perfil	3	8,82	Indeterminado	9	10,34
Misto	5	14,70			

Fonte: Jaiswal, A. K., Singh, N., & Ahuja, N. J. (2017). Learning styles based adaptive intelligent tutoring systems: Document analysis of articles published between 2001. and 2016. *International Journal of Cognitive Research in Science, Engineering and Education:(IJCRSEE)*, 5(2), 83-98.

Uma análise crítica, de 78 pesquisas realizadas de 2000 a 2013, foi apresentada por Al-Azawei e Badii (2014). Os principais métodos de análise de dados usados para a detecção implícita baseada em dados foram Redes Bayesianas, Árvores de Decisão, Modelo Oculto de Markov e Redes Neurais. A maior parte dos estudos usou sistemas personalizados. Dos estudos que criaram complementos para ferramentas *open source*, o AVA mais utilizado foi o Moodle. O modelo de estilo de aprendizagem mais usado foi o Felder-Silverman e os pontos

apontados para isso foram: (i) classifica os estudantes em 16 estilos; (ii) a medida tem validade e confiabilidade; (iii) mais consistente com as preferências dos aprendizes; e (iv) adequação e aplicação prática do modelo para a potencialização da aprendizagem por meio da tecnologia. A conclusão foi que a área ainda é emergente. As amostras usadas nos estudos revisados foram muito pequenas e em contextos muito específicos, diminuindo a generalidade dos resultados. Os principais desafios da detecção automática de estilos de aprendizagem para SEAHs apontados foram a escolha de um modelo teórico de estilo de aprendizagem que privilegie a adaptabilidade do sistema, identificar a abordagem mais adequada para deduzir os estilos e lidar com o *cold-start* (fato dos usuários não terem qualquer informação no sistema antes de entrarem nele) (Al-Azawei & Badii, 2014).

É fundamental que os modelos de estilos de aprendizagem tenham uma base teórica consistente e de natureza psicológica. Foram identificadas algumas tentativas de profissionais da área de computação de gerarem modelos de estilos de aprendizagem, o que não é aconselhável. Muitos modelos, apesar de apresentarem boa base psicológica, são muito complexos para serem aplicados. Os instrumentos precisam ter excelentes validade e confiabilidade, uma vez que a detecção é baseada neles (Al-Azawei & Badii, 2014). A detecção explícita foi quase que completamente substituída pela detecção implícita nos últimos anos. Cada vez mais esse modelo é usado por causa de sua capacidade de deduzir e automaticamente atualizar as preferências dos estudantes. No entanto, ainda são poucos e inovadores os modelos dinâmicos de detecção, que são necessários, uma vez que os estilos de aprendizagem podem variar com o tempo (Al-Azawei & Badii, 2014).

O problema do *cold-start* é muito presente nos estudos analisados por criarem sistemas e cursos do zero e todos os estudantes entrarem na plataforma e no curso como primeira interação. Ou seja, por não haver informações anteriores sobre esses estudantes, não é possível definir um estilo no momento da entrada no curso. Uma das soluções é aplicar um

questionário de estilos de aprendizagem assim que o estudante entra na plataforma. Outra solução é atribuir um estilo *default* aos estudantes e atualizar essa informação à medida que ele se comporta no sistema (Al-Azawei & Badii, 2014). Outro ponto relevante é a dependência da detecção automática de escalas de estilos e aprendizagem. Todos os estudos apontam como problema a validade psicométrica dos instrumentos, no entanto, a única forma usada para avaliar a precisão da detecção automática foi a comparação com esses mesmos instrumentos (Al-Azawei & Badii, 2014).

Para complementar as revisões de literatura citadas, foram analisados os trabalhos publicados de 2015 a 2018 sobre detecção automática de estilos de aprendizagem. Os trabalhos foram buscados na plataforma Google Scholar usando, em conjunto, os termos “*automatic detection*” e “*learning style*”. A busca retornou 266 resultados, dos quais foram selecionados os que criaram ou aprimoraram uma ferramenta de identificação automática de estilos de aprendizagem. Os 21 artigos selecionados foram analisados quanto ao modelo de estilo de aprendizagem, à abordagem, ao método de análise de dados, ao AVA e à quantidade de cursos para os quais a ferramenta foi desenvolvida e a quantidade de estudantes usados para a criação das ferramentas (Apêndice B). Comparando com as revisões anteriores, percebe-se um aumento na quantidade de estudantes considerados para criar as ferramentas. Apenas três estudos usaram amostras menores do que 50 e sete (33%) incluíram amostras menores do que 100. Por outro lado, ainda são muito escassas as ferramentas desenvolvidas para mais de um curso (14%). Quanto aos modelos de estilos de aprendizagem, o de Felder-Silverman é amplamente mais utilizado (71%) do que os outros e não foram encontrados estudos que utilizaram mais de um modelo.

A maior mudança em relação ao descrito nas revisões anteriores foram as abordagens. Ao contrário do que foi apontado por Feldman *et al.* (2015), houve uma predominância pelo uso da abordagem baseada na literatura, de 2015 a 2018. Uma inovação metodológica, não

identificada pelas revisões anteriores, foi o uso das abordagens baseada em dados e em literatura no mesmo estudo para melhorar de forma dinâmica a validade do estilo de aprendizagem identificado (Arslan, 2018; Sahid, Nugroho & Santosa, 2017). Além disso, a identificação dinâmica dos estilos de aprendizagem esteve presente em 28% dos estudos. O *software* mais usado para a análise de dados foi o WEKA.

Os resultados da revisão realizada para esta Dissertação apontam que as pesquisas na área têm desenvolvido soluções para aumentar a validade das medidas de estilos de aprendizagem aferidos, com o uso de técnicas de natureza dinâmica e de diversos algoritmos e comparando a validade deles, como em Bernard, Chang, Popescu e Graf (2016), e privilegiando o uso de abordagens baseadas em literatura, uma vez que elas são apontadas como tendo maior validade (Graf, 2007). Esses são métodos que demandam não só mais trabalho dos pesquisadores, como um conhecimento aprofundado de psicologia e do curso, para serem implementados. Junto a isso, não houve avanço quanto à generalidade das ferramentas, uma vez que quase todas elas podem ser aplicadas para apenas um curso. A exceção foi o estudo de Rajper, Saikh, Saikh e Mallah (2016). Os autores desenvolveram uma ferramenta que não fosse dependente de um ou mais cursos e sim aplicável a um grupo deles. No entanto, a ferramenta, após seu desenvolvimento, foi testada em apenas um curso com 20 estudantes.

3.3 Criação de Ferramenta de Detecção Automática de Estilos de Aprendizagem

O passo seguinte foi a construção da ferramenta de detecção automática dos estilos de aprendizagem aferidos pela EEA-Ead_C. Para isso, foram criados bancos de dados com *footprints* dos estudantes no AVA e estes foram transformados em indicadores. Em seguida, os indicadores foram selecionados, tratados e agrupados. Após o tratamento, eles foram relacionados com os estilos de aprendizagem autodeclarados pelos estudantes nas respostas ao questionário. Por fim, foi realizada uma tentativa de criar um algoritmo para prever os estilos de aprendizagem autodeclarados usando os indicadores criados.

O objetivo da criação da ferramenta é que ela seja aplicada a uma plataforma e não a um curso específico. Para que isso fosse possível, foram adotadas duas premissas:

1. Apesar dos comportamentos ligados à aprendizagem variarem por curso, a frequência relativa do estudante em relação a seus colegas varia pouco.
2. Os estilos de aprendizagem se mantêm constantes no tempo.

As implicações práticas da primeira premissa são que (i) todos os indicadores foram usados na forma de escores z para permitir que estudantes de cursos diferentes fossem comparados; (ii) foram analisados apenas os estudantes que foram aprovados nos cursos. Se os estudantes que não fossem aprovados fossem analisados junto com os aprovados, isso geraria uma distorção na frequência dos comportamentos, uma vez que eles teriam menos oportunidades de apresentar os comportamentos. A relevância de se detectar estilos de aprendizagem dessa forma é que, uma vez que o estudante tivesse finalizado um curso, ele já teria seu estilo identificado e essa informação seria levada para cursos subsequentes. Como a definição utilizada para estilo de aprendizagem é que ele é um padrão comportamental frequente e expresso em diversos contextos, também foi analisada a constância dos comportamentos com base nos indicadores levantados.

Para as análises de dados foram usados os *softwares* Excel 2013 para transformações, o SPSS 23 para transformações, análises de comparação de médias e análises fatoriais e o WEKA para a criação dos algoritmos de detecção de estilos. O WEKA é um *software* criado para facilitar o uso de técnicas de *machine learning*. Ele tem uma interface simples de usar e que faz o pré-processamento e a análise dos dados (Witten *et al.*, 1999). Na análise da literatura de detecção automática de estilos de aprendizagem, o WEKA foi o *software* mais usado.

3.3.1 Levantamento de Indicadores

Com base nas informações da estrutura relacional do banco de dados (Apêndice C) e no questionário construído, foram definidos os comportamentos relevantes para a detecção automática dos estilos de aprendizagem. Os 39 comportamentos definidos foram categorizados em relação à categoria teórica de estilo de aprendizagem a que fazem parte (regulação ou interação) e a qual tipo de informação em banco de dados se referem (ordem, tempo, quantidade de razão), conforme tabela no Apêndice F. Os recursos de cada curso também foram ligados a uma de três categorias (avaliação, recurso e interação), conforme Apêndice G.

Para serem transformados em indicadores, as informações referentes aos comportamentos poderiam ser usadas de forma bruta, como média ou como desvio-padrão. Além disso, elas também poderiam ser comparadas com outros recursos de sua categoria. Por exemplo, para postagens em fóruns, poderiam ser criados os indicadores de quantidade bruta de postagens em fóruns, média de postagens por fórum, desvio-padrão de postagens por fórum e relação entre postagens em fóruns e postagens em discussões. Sendo assim, os 39 comportamentos poderiam ser transformados em aproximadamente 150 indicadores. No entanto, há três questões importantes ao manipular bancos de dados. A primeira delas é a disponibilidade da informação, que precisa ser checada caso a caso. Por exemplo, em análises preliminares do banco, esperava-se que a informação sobre a conclusão do curso estivesse na tabela de conclusões dos cursos. No entanto, essa tabela estava vazia quando foi analisada e foi necessário encontrar outra estratégia para identificar quem finalizou ou não o curso. A segunda é a confiabilidade da informação. No caso dos indicadores de ordem, onde era necessário identificar a ordem natural do curso para checar, por exemplo, se o estudante a havia seguido, a única informação encontrada no banco foi a data de criação do recurso. Ordenar recursos por sua data de criação não necessariamente coincidiria com a ordem na

qual o curso foi montado. Outro exemplo disso são os indicadores de duração. O banco de dados disponibilizado traz apenas a informação da data e do horário onde o comportamento ocorreu. A duração deveria ser inferida comparando um determinado comportamento com o que foi realizado logo em seguida. No entanto, esse tempo não necessariamente teria sido usado para realizar a atividade em questão. Dessa forma, os indicadores de duração não seriam confiáveis. A última questão é quanto esforço é necessário para obter a informação. Considerando que vários cursos serão analisados de uma vez e que o objetivo deste trabalho é criar uma solução genérica para identificar os estilos de aprendizagem, seria contraditório usar informações que demandam muito esforço para serem obtidas. Por exemplo, o banco de dados não liga os recursos disponibilizados à avaliação. Dessa forma, para identificar se o estudante acessou todos os recursos relevantes para realizar uma avaliação antes de realizá-la, seria necessário analisar cada recurso e cada prova e associá-los.

Considerando as restrições impostas para o levantamento de indicadores, optou-se por indicadores que atendessem plenamente aos critérios de disponibilidade, confiabilidade e esforço. Foram extraídas 43 informações do banco de dados (Apêndice H). Dessas informações, três são básicas, que servem para identificar os usuários e os cursos. Outras 11 são informações de comportamento dos estudantes em nível de curso. As outras 29 são comportamentos do próprio estudante no curso. As informações extraídas foram transformadas em 25 indicadores (Apêndice I), que foram usados para as análises posteriores. Para que estudantes de diferentes cursos pudessem ser comparados, foi calculado o escore z de cada um deles em cada um dos cursos.

3.3.2 Amostra de Cursos

O banco conta com 305 cursos. Destes, alguns são comunidades de prática, outros repositórios e outros são cursos com atividades instrucionais. Para a detecção dos estilos de aprendizagem, foram somente escolhidos os cursos com tais atividades ($N = 92$). Destes,

foram retirados os cursos com menos de 20 aprovados ou que não tivessem interação. Com esse novo filtro, a quantidade de cursos baixou para 36, compostos por 2.367 registros de 2.094 estudantes aprovados. Os cursos são referentes a competências técnicas específicas do órgão e a competências técnicas genéricas, como Excel avançado e tomada de decisão. A lista dos cursos não é apresentada para não identificar o órgão onde foi realizada a pesquisa. Para definir um estudante como aprovado, foram adotados como critério a realização de 70% ou mais das atividades avaliativas e do atingimento de média final igual ou superior a 70%.

3.3.3 *Análise de Indicadores*

Para definir se um indicador seria usado ou não, foram analisados sua variabilidade, quantidade de dados omissos, variação para o mesmo estudante e relação com os estilos de aprendizagem autodeclarados. Para usar os critérios descritos, os indicadores foram analisados para cada um dos 36 cursos. Foi observado que alguns cursos limitavam o comportamento dos estudantes. Por exemplo, para alguns deles, a quantidade de tentativas para os *quizzes* era igual a um. Sendo assim, todos os aprovados naquele curso haviam tentado realizar os *quizzes* apenas uma vez. Sem variabilidade, os escores z foram iguais a zero. Esses casos foram categorizados, mas, para o uso dos indicadores como variáveis contínuas, eles foram mantidos como zero. Para os casos onde o comportamento típico absoluto era zero, o dado foi classificado como omissos.

Foi realizada a análise descritiva dos indicadores e poucos deles se aproximavam de uma distribuição normal (Tarefas respondidas; Tempo restante para a entrega; Tempo após divulgação; e Fóruns X Leitura de Fóruns). Para esses indicadores, os casos extremos foram reclassificados para 1,96 e -1,96, que são escores referentes a dois desvios-padrões para cima ou para baixo. Isso não foi feito nos outros indicadores porque suas distribuições se assemelham a uma distribuição de cauda longa, onde há muitos casos extremos e muito extremos negativos ou positivos. Todos os indicadores foram normalizados pelo

procedimento de ranqueamento fracional, que é indicado para distribuições com grande assimetria (Leydesdorff, Zhou & Bornmann, 2013).

Para testar a hipótese de que o comportamento do estudante comparado ao de seus colegas se mantém relativamente constante em diferentes cursos, foram analisados os casos de estudantes que foram aprovados em mais de um dos 36 cursos. Foram analisados 219 estudantes em 492 incidências, onde a quantidade de cursos por estudante variou de 2 a 5. Foram calculados a média e o desvio-padrão do estudante para cada um dos indicadores. Os casos foram categorizados na quantidade de (i) desvios-padrões menores do que 0,5; (ii) desvios-padrões entre 0,5 e 1; (iii) desvios-padrões entre 1 e 1,5; (iv) desvios-padrões entre 1,5 e 2; e (v) desvios-padrões maiores do que 2. Foram calculados a média e o desvio-padrão do percentual de indicadores em cada grupo por estudante (Figura 12). Na média, os estudantes têm 76% dos indicadores com até um desvio-padrão, no entanto, o desvio-padrão dessa média é muito alto (17%). Essa informação mostra de forma conjunta a variabilidade em relação aos indicadores dos estudantes e dos próprios indicadores.

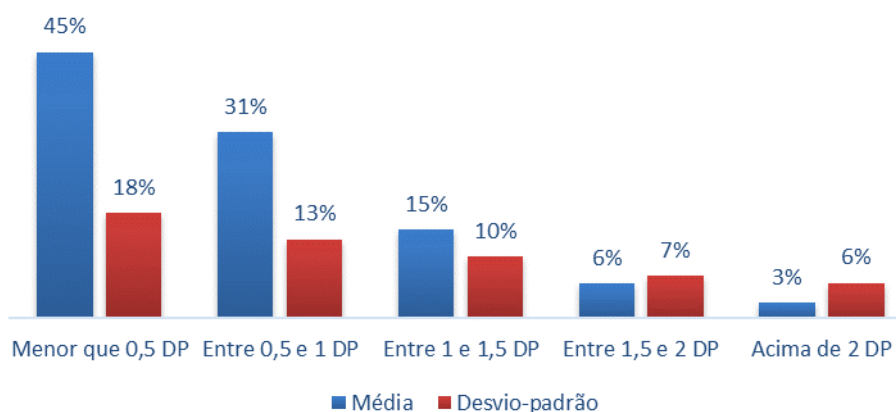


Figura 12. Proporção média de indicadores por categoria de desvio-padrão.

Considerando que o desvio-padrão das categorias foi muito alto em relação à média, foram investigados separadamente os estudantes e os indicadores. Foi identificado o percentual de indicadores que variaram mais de um desvio-padrão para cada estudante (Figura

13). A maior parte dos estudantes (69%) tem entre 10 e 50% dos indicadores variando mais de um desvio-padrão. Apenas 6% dos estudantes têm mais de 50% dos indicadores variando mais de um desvio-padrão.

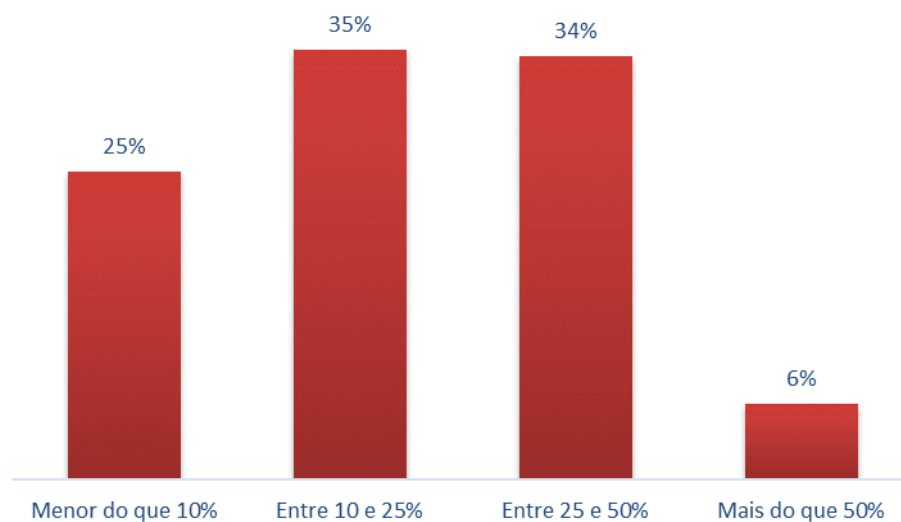


Figura 13. Distribuição de estudantes em relação ao percentual de indicadores que variaram mais de um desvio-padrão.

Os indicadores também foram analisados quanto ao percentual de incidências de desvios-padrões superiores a 1. Houve grande variação na constância quanto aos indicadores (Figura 14), sendo que o que variou menos foi o percentual de tarefas respondidas (7%) e o que mais variou foi a quantidade de discussões criadas (37%).

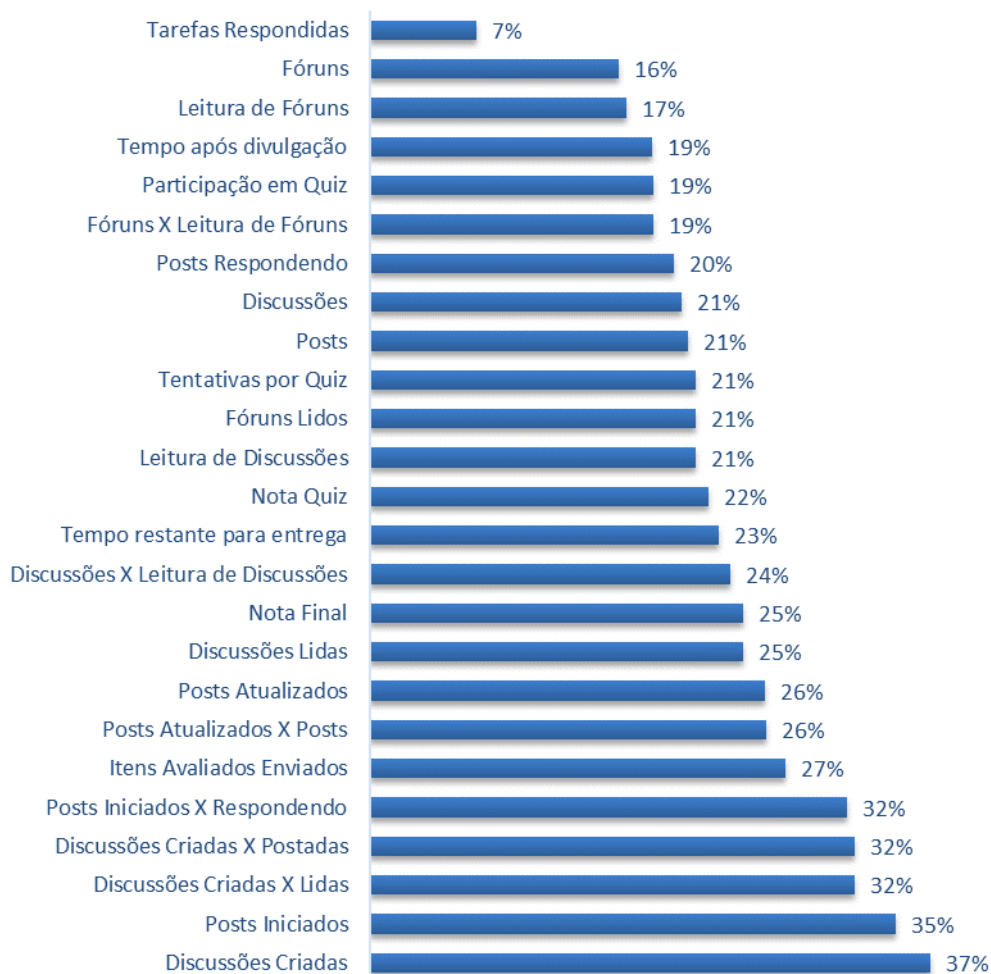


Figura 14. Percentual de variações superiores a um desvio-padrão por indicador.

Os indicadores normalizados foram analisados por meio de análise fatorial exploratória. Ao analisar todos os indicadores em conjunto, correlações entre os indicadores de interação e de avaliação foram muito baixas. Não foi possível encontrar uma solução fatorial que incluísse todos os indicadores. Sendo assim, os indicadores foram analisados separadamente. Para os indicadores de avaliação, não foi possível encontrar uma solução fatorial com boa validade estatística. Na análise feita separadamente para os indicadores de interação, foi usado um terço do banco de dados, escolhido aleatoriamente, e foram escolhidos apenas os estudantes que não tiveram dados omissos para qualquer indicador. Os 484 casos foram analisados usando o método de fatoração por quadrados mínimos não ponderados e a rotação usada foi a promax. O KMO foi de 0,81, o que é considerado ótimo

(Hutcheson e Sofroniou, 1999). A matriz de correlações mostrou mais de 50% de correlações acima de 0,3. Analisando a matriz de correlações, encontrou-se que as correlações entre os indicadores de interação foram consideravelmente altas (13% acima de 0,8). As comunalidades foram altas (Apêndice J), sendo 15 das 17 acima de 0,8. Ainda havendo uma possível redundância entre os indicadores de interação, todos foram mantidos. O critério dos autovalores superiores a 1 apontou quatro fatores com variância explicada de 83% (Tabela 12). O *scree plot* também apontou quatro fatores (Apêndice K).

Tabela 12

Variância total explicada da análise fatorial exploratória dos indicadores de interação

Fator	Autovalores iniciais			Somadas de extração de carregamentos ao quadrado			Somadas de rotação de carregamentos ao quadrado
	Total	% de variância	% cumulativa	Total	% de variância	% cumulativa	Total
1	7,645	44,970	44,970	7,513	44,195	44,195	5,996
2	3,386	19,920	64,890	3,144	18,493	62,688	6,157
3	2,129	12,523	77,413	1,987	11,686	74,374	4,144
4	1,568	9,226	86,639	1,480	8,704	83,077	2,900

A estrutura com quatro fatores apresentou boas consistências teórica e estatística, com cargas compartilhadas com diferença em módulo superiores a 0,1 (Tabela 13). O Fator 1 é composto por cinco indicadores e foi denominado de Interação Proativa. Fazem parte dele os indicadores de criação de discussão e de postagens iniciadas. O Fator 2 é composto por quatro indicadores e foi denominado de Interação Ativa. Fazem parte dele os indicadores brutos de postagens, discussões, fóruns e postagens de resposta. O Fator 3 é composto por seis indicadores e foi denominado de Interação Passiva. Fazem parte dele todos os indicadores que envolvem acesso à interação de terceiros. O Fator 4 é composto por dois indicadores e foi denominado de Atualização. Fazem parte dele os indicadores de atualização de postagens. A consistência interna dos fatores, medida pelo alfa de Cronbach, varia de 0,67 a 0,99 (Tabela 14).

Tabela 13
Matriz de padrão da análise fatorial exploratória dos indicadores de interação

	Fator			
	1	2	3	4
Posts		0,931		
Posts Iniciados	0,866			
Posts Respondendo		1,041		
Posts Iniciados X Respondendo	0,923			
Discussões		0,955		
Fóruns		0,958		
Leitura de Fóruns		0,319	0,682	
Leitura de Discussões			0,792	
Discussões X Leitura de Discussões		0,315	-0,865	
Fóruns X Leitura de Fóruns		0,355	-0,793	
Fóruns Lidos		0,351	0,532	
Discussões Lidas			0,82	
Discussões Criadas	0,936			
Discussões Criadas X Postadas	1,012			
Discussões Criadas X Lidas	1,006			
Posts Atualizados				1,001
Posts Atualizados X Posts				0,936

Tabela 14
Nomes e consistência interna dos fatores de indicadores de interação

	Nome	N Indicadores	Alfa de Cronbach
Fator 1	Interação Proativa	5	0,98
Fator 2	Interação Ativa	4	0,99
Fator 3	Interação Passiva	6	0,67
Fator 4	Atualização	2	0,95

Foi calculada a média ponderada pela carga fatorial dos indicadores de cada fator para as análises subsequentes. Os fatores foram analisados em relação a sua constância para os estudantes que realizaram mais de um curso (Figura 15). O fator mais constante foi o de Interação Ativa, com apenas 17% das incidências com variação superior a um desvio-padrão. Enquanto que o fator menos constante foi o de Interação Proativa, com 28% de incidências superiores a um desvio-padrão.

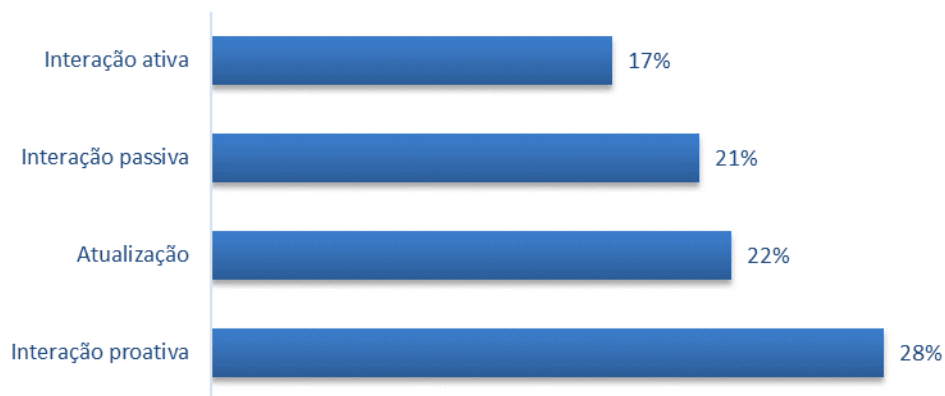


Figura 15. Percentual de variações superiores a um desvio-padrão por fator.

3.3.4 Categorização dos Estilos de Aprendizagem Autodeclarados

Os algoritmos de predição utilizados classificam os indivíduos em uma determinada categoria da variável dependente de acordo com as variáveis independentes. Sendo assim, foi necessário categorizar os estudantes de acordo com seus estilos de aprendizagem autodeclarados. Eles foram divididos em 4 categorias usando a análise de clusters *k-means*. As categorias usadas foram (i) baixas organização e potencialização da aprendizagem e abordagem superficial; (ii) baixas organização e potencialização da aprendizagem e abordagem profunda; (iii) altas organização e potencialização da aprendizagem e abordagem superficial; e (iv) altas organização e potencialização da aprendizagem e abordagem profunda. A categoria com mais casos foi a primeira (500) e com menos casos foi a terceira (170), conforme Tabela 15.

Tabela 15
Quantidade de casos por cluster

	Nome	N
Cluster 1	Baixas organização e potencialização da aprendizagem e abordagem superficial	500
Cluster 2	Baixas organização e potencialização da aprendizagem e abordagem profunda	383
Cluster 3	Altas organização e potencialização da aprendizagem e abordagem superficial	170
Cluster 4	Altas organização e potencialização da aprendizagem e alta abordagem profunda	423
Não classificado	-	286

Uma vez que as bases teóricas usadas para construir a medida apontam que os estilos de aprendizagem são um contínuo (Entwistle, 1981; Vermunt, 1994), a categorização dos indivíduos é artificial. Isso se reflete na distribuição dos dados, onde a distância entre o caso e dois centroides é muito similar. Sendo assim, a análise foi feita em dois passos, classificando cada um dos fatores em dois *clusters* (Figura 16). Os casos que estavam muito próximos do centroide de outro *cluster* não foram categorizados. O critério para exclusão foi a diferença entre as distâncias do caso para o centroide ao qual foi atribuído e o segundo centroide mais próximo. Quando a distância foi inferior a 30% do desvio-padrão da média ponderada do fator, o caso não foi classificado.

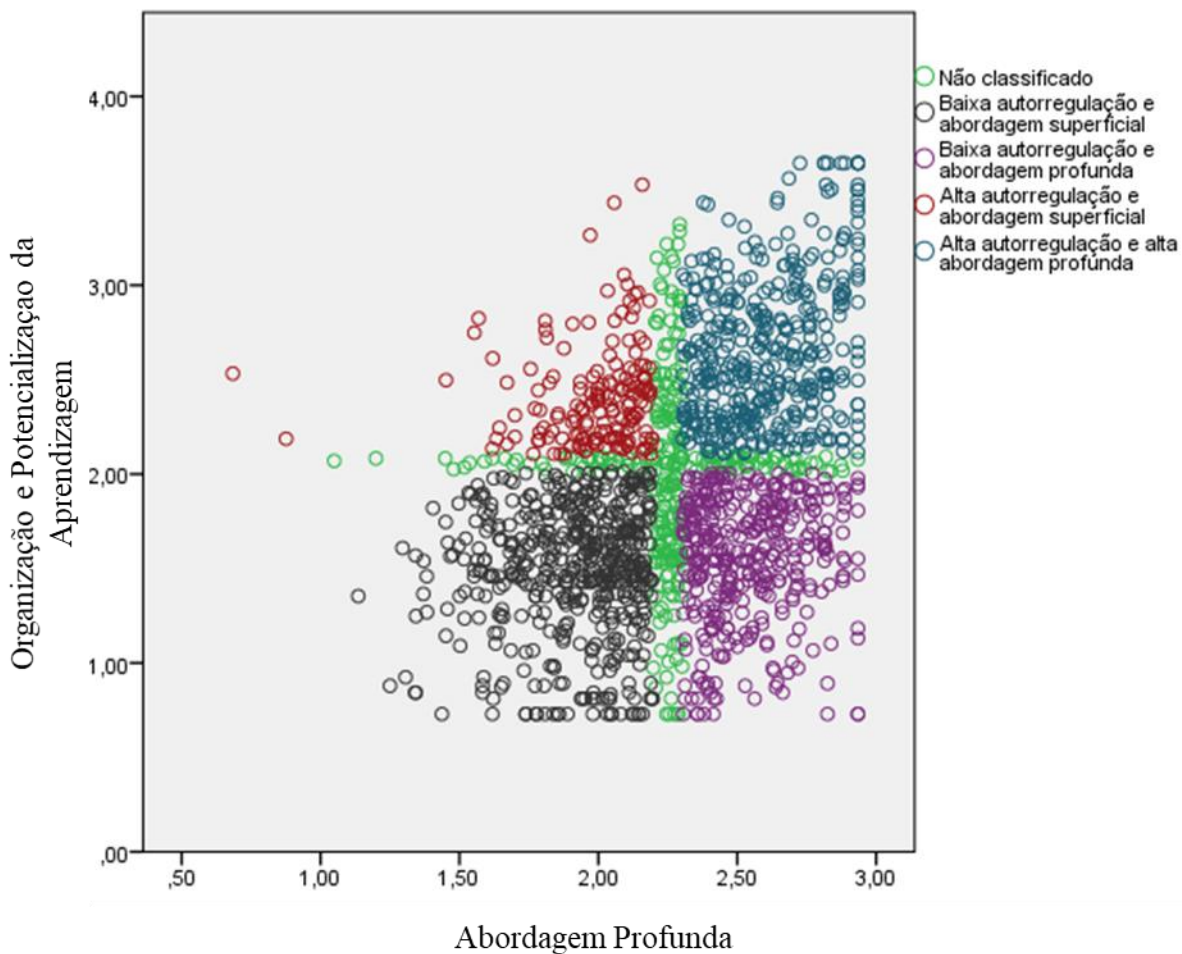


Figura 16. Gráfico de dispersão por fator de estilo de aprendizagem autodeclarado com classificação dos casos.

Além disso, a fim de testar a predição dos indicadores em relação a cada um dos fatores da escala, eles também foram separadamente submetidos à análise de *clusters*. Os estudantes foram colocados nas categorias (i) baixa; (ii) média; (iii) alta; e (iv) não categorizado (Tabela 16). O critério usado para que um estudante não fosse categorizado foi o mesmo usado na análise de *cluster* anterior.

Tabela 16
Clusters de cada um dos fatores de estilo de aprendizagem autodeclarado

Clusters por fatores	
Fator 2 - Organização e Potencialização da Aprendizagem	
<i>Baixa</i>	555
<i>Média</i>	570
<i>Alta</i>	309
<i>Não classificada</i>	328
Fator 1 - Abordagem profunda	
<i>Baixa</i>	276
<i>Média</i>	600
<i>Alta</i>	523
<i>Não classificada</i>	361

3.3.5 Relação entre Indicadores e Estilos Autodeclarados

Ao selecionar os estudantes aprovados em um dos 36 cursos e que responderam ao questionário, a amostra era de 251 casos. Desses casos, 55 são de estudantes que realizaram mais de um curso. Para esses casos, foi adotada a média de cada indicador e cada fator do estudante. Para os casos onde o desvio padrão foi maior do que 1, o escore no indicador foi classificado como omissos. Foram excluídos os estudantes que tiveram o desvio-padrão maior do que 1 em 4 ou mais dos outros fatores, pois não foi possível identificar o estilo de aprendizagem deles por meio de indicadores. Foram eliminados 3 casos por esse critério. A amostra final contou com 212 casos.

Para explorar a relação entre os indicadores e os estilos autodeclarados, foram feitas correlações usando a correlação de Pearson. Foram realizadas análises individuais para cada indicador e cada fator e as condições da variável critério. Os resultados estão resumidos na Tabela 17. A maior parte dos indicadores tem pouca ou nenhuma relação com as condições do estilo de aprendizagem autodeclarado, sendo os indicadores de interação os que mais apresentam correlações significativas. Quanto aos fatores observados de interação, o Fator Interação Proativa e Interação Ativa possuem boas correlações com os estilos autodeclarados.

Tabela 17
Correlações de Pearson entre os indicadores normalizados, os fatores de interação e as variáveis dependentes

	4 Clusters - 2 Fatores	3 Clusters - Fator 1	3 Clusters - Fator 2
Interação Proativa	0,238**	0,174*	0,158*
Interação Ativa	0,288**	0,115	0,332**
Interação Passiva	0,081	-0,086	0,152
Atualização	0,094	0,118	0,032
Tarefas Respondidas	0,057	0,064	0,164
Tempo restante para entrega	-0,145	-0,161	-0,039
Tempo após divulgação	-0,203	-0,127	-0,047
Participação em Quiz	0,079	0,183*	0,036
Tentativas por Quiz	0,133	0,039	0
Nota Quiz	0,064	-0,011	0,024
Itens Avaliados Enviados	0,102	0,155	0,049
Nota Final	0,091	0,008	0,066
Posts	0,293**	0,117	0,349**
Posts Iniciados	0,280**	0,13	0,225**
Posts Respondendo	0,235**	0,097	0,294**
Posts Iniciados X Respondendo	0,196*	0,081	0,119
Discussões	0,302**	0,126	0,331**
Fóruns	0,292**	0,105	0,310**
Leitura de Fóruns	0,197**	0,052	0,224**
Leitura de Discussões	0,11	-0,05	0,213**
Discussões X Leitura de Discussões	0,180*	0,186*	0,113
Fóruns X Leitura de Fóruns	0,079	0,073	0,113
Fóruns Lidos	0,143	0,01	0,186*
Discussões Lidas	,161*	-0,05	0,239**
Discussões Criadas	0,251**	,206**	0,177*
Discussões Criadas X Postadas	0,196**	0,180*	0,115
Discussões Criadas X Lidas	0,192*	0,157*	0,123
Posts Atualizados	0,111	0,133	0,045
Posts Atualizados X Posts	0,049	0,1	0,038

** A correlação é significativa no nível 0,01 (bilateral).

* A correlação é significativa no nível 0,05 (bilateral).

3.3.6 Teste da Detecção Automática de Estilos de Aprendizagem

Os indicadores e os fatores foram usados para criar o algoritmo de predição de estilo de aprendizagem. Foram usadas as técnicas *Naïve Bayes*, *Random Tree*, *Random Forest* e J48 usando 70% do banco de dados para treinamento (N = 148) e 30% para teste (N = 64). No entanto, para nenhuma das condições foi encontrado um poder de predição superior a 55%. Considerando a elevada taxa de erro do algoritmo criado, os estilos de aprendizagem autodeclarados não foram preditos pelos indicadores.

3.4 Comparação de Medidas

Uma vez que não foi possível inferir os estilos de aprendizagem autodeclarados a partir dos estilos observados, as duas medidas foram testadas em relação a suas validades de critério. A variável escolhida para este teste foi a evasão. O critério adotado para classificar um estudante como evadido foi que ele tivesse realizado pelo menos uma das atividades do curso e menos de 70% das atividades avaliativas. Os estudantes que realizaram mais de 70% das atividades avaliativas, mas obtivessem notas médias inferiores a 70% foram classificados como reprovados. Ao todo, foram 3.883 incidências de 3.405 estudantes referente aos 36 cursos que estão sendo analisados. Das incidências, 2.367 são de aprovações, 1.390 de evasões e 126 de reprovações. Os estudantes foram classificados em dois grupos, os que evadiram pelo menos um curso e os que não evadiram nenhum curso.

Inicialmente, os estilos de aprendizagem autodeclarados foram testados em relação à evasão. Da amostra de 3.405 estudantes, 311 responderam ao questionário. Destes, 208 foram classificados como não tendo evadido nenhum curso e 102 como tendo evadido pelo menos um curso. Foi criada uma amostra balanceada e essa amostra foi testada no *software* SPSS por meio de regressão logística. Foram usadas como variáveis independentes os escores fatoriais do Fator Organização e Potencialização da Aprendizagem e do Fator Abordagem Profunda e a quantidade de cursos dos quais o estudante participou. O teste de qui-quadrado apontou que

os coeficientes das variáveis preditoras não são diferentes de zero (Tabela 18). Os R^2 de Cox e Snell e de Nagelkerke foram iguais a zero, o que indica que não há variância explicada pelo modelo. A classificação dos casos usando o modelo obteve um percentual de acerto de 51,2%, sendo que todos os casos foram classificados como não evadidos (Tabela 19).

Tabela 18

Teste de Omnibus da regressão logística dos dois fatores de estilo de aprendizagem autodeclarado e a quantidade de cursos realizados em relação à evasão

		Qui-quadrado	gl	Sig.
Passo 1	Passo	0,023	2	0,989
	Bloco	0,023	2	0,989
	Modelo	0,023	2	0,989

Tabela 19

Classificação de casos da regressão logística dos dois fatores de estilo de aprendizagem autodeclarado e a quantidade de cursos realizados em relação à evasão

Observado			Predito		
			Eva_bin		Porcentagem correta
			Não	Sim	
Passo 0	Eva_bin	Não	108	0	100,0
		Sim	103	0	0,0
Porcentagem global					51,2

A mesma amostra balanceada foi testada usando o *software* WEKA. Foi usada a técnica *Random Committee*, que é um tipo de Comitê de Classificadores. Esse método constrói um *output* com resultados de vários classificadores para melhorar o desempenho da classificação (Witten e Frank, 2016). Usando 79% da amostra para treinamento ($N = 167$) do algoritmo e 21% para teste ($N = 44$), a taxa de acerto do modelo foi de 63% (Tabela 20).

Tabela 20

Classificação de casos usando Random Committee para os dois fatores de estilo de aprendizagem autodeclarado e a quantidade de cursos realizados em relação à evasão

Observado		Predito		
		Evasão		Porcentagem correta
		Não	Sim	
Evasão	Não	17	3	51,5
	Sim	16	8	72,7
Porcentagem global				63,1

Inicialmente, os estilos de aprendizagem observados foram testados em relação à evasão. Da amostra de 3.405 estudantes, 2.094 foram aprovados em pelo menos um curso e foram classificados em relação a seus indicadores. Destes, 1.984 foram classificados como não tendo evadido nenhum curso e 110 como tendo evadido pelo menos um curso. Foi criada uma amostra balanceada e essa amostra foi testada no *software* SPSS por meio de regressão logística. Foram usadas como variáveis independentes os escores fatoriais dos fatores Interação Proativa, Interação Ativa, Interação Passiva e Atualização de Interação e a quantidade de cursos dos quais o estudante participou. O teste de qui-quadrado apontou que os coeficientes das variáveis predictoras não são diferentes de zero (Tabela 21). Os R^2 de Cox e Snell e de Nagelkerke foram iguais a, respectivamente, 0,01 e 0,013, o que indica que a variância explicada no modelo é praticamente igual a zero. A classificação dos casos usando o modelo obteve um percentual de acerto de 53,5%, sendo que a previsão referente ao grupo de evadidos obteve um percentual de acerto significativamente abaixo do aleatório (29,4%) (Tabela 22).

Tabela 21

Teste de Omnibus da regressão logística dos quatro fatores de estilo de aprendizagem observado e a quantidade de cursos realizados em relação à evasão

		Qui-quadrado	gl	Sig.
Passo 1	Passo	2,110	5	,834
	Bloco	2,110	5	,834
	Modelo	2,110	5	,834

Tabela 22

Classificação de casos da regressão logística dos quatro fatores de estilo de aprendizagem observado e a quantidade de cursos realizados em relação à evasão

Observado			Predito		Porcentagem correta
			Evasão		
			Não	Sim	
Passo 1	Evasão	Não	84	27	75,7
		Sim	72	30	29,4
Porcentagem global					53,5

A mesma amostra balanceada foi testada usando o *software* WEKA. Foi usada a técnica *Locally Weighed Learning*. Esse é um método não paramétrico de aproximação de funções, onde várias funções locais são criadas (Englert, 2012). Usando 77% da amostra para treinamento (N = 164) do algoritmo e 23% para teste (N = 49), a taxa de acerto do modelo foi de 69% (Tabela 23).

Tabela 23

Classificação de casos usando Locally Weighed Learning para os dois fatores de estilo de aprendizagem autodeclarado e a quantidade de cursos realizados em relação à evasão

Observado			Predito		Porcentagem correta
			Evasão		
			Não	Sim	
Evasão	Não		24	4	64,9
	Sim		13	11	73,3
Porcentagem global					68,8

3.5 Discussão

A nova medida criada de estilos de aprendizagem baseada em frequência de comportamentos (EEA-Ead_C) tem bons indicadores de validade estatística e o agrupamento dos fatores está de acordo com uma teoria válida do construto (Entwistle, 1981). No entanto, sua estrutura diferiu consideravelmente da estrutura hipotetizada pelos pesquisadores e pelos

juízes e pela estrutura do questionário que foi usado como base (Jesus, 2014). Um dos possíveis motivos para isso é que a percepção do indivíduo sobre suas frequências comportamentais nos cursos difere consideravelmente de sua preferência. Sendo assim, o que é preferido não se agrupa da mesma forma do que se percebe que é realizado. É necessário que os itens sejam revistos, dessa vez levando em conta a estrutura encontrada, e que a medida seja validada para outros contextos. O novo agrupamento trazido pela medida teve um impacto significativo ao relacioná-la com os indicadores comportamentais, uma vez que os agrupamentos encontrados dos indicadores comportamentais, especificamente os de interação, aproximaram-se do agrupamento hipotetizado pelos pesquisadores e juízes.

As pesquisas em detecção automática de estilos de aprendizagem fazem a detecção para um curso apenas ou para um número muito pequeno deles. Além disso, as informações usadas para detectar os estilos estão muito ligadas ao conteúdo do curso, se apoiam em modelos que também são ligados a ele, como o Felder-Silverman (Felder & Silverman, 1988), e usam a detecção para personalizar a apresentação de conteúdo (Al-Azawei & Badii, 2014; Jaiswal *et al.*, 2017; Feldman *et al.*, 2015; e Oliveira *et al.*, 2018). No entanto, os próprios criadores das teorias desencorajam a apresentação de conteúdo baseada em estilos (Felder & Spurling, 2005). A identificação de padrões comportamentais relativamente estáveis de aprendizagem pode trazer muitos benefícios além da apresentação do conteúdo. Nesta dissertação buscou-se aliar uma medida que fosse baseada em comportamentos referentes a estilos de aprendizagem que não estivessem ligadas ao conteúdo para que elas pudessem ser generalizadas e usadas em mais de um curso. Buscou-se criar um algoritmo que servisse para um AVA e não apenas para um curso. Para a organização onde o trabalho foi realizado, não foi possível desenvolver esse algoritmo de detecção automática dos estilos de aprendizagem com um poder de predição aceitável. Alguns dos possíveis motivos para isso são: (i) a diferença de tempo entre a aplicação do questionário e a realização dos

comportamentos, que foi na maioria dos casos, superior a um ano; (ii) a diferença entre a percepção do indivíduo em relação a sua frequência comportamental e sua frequência comportamental; (iii) o uso apenas de indicadores de avaliação e de interação; (iv) a complexidade dos construtos de estilo de aprendizagem autodeclarado; (v) a baixa quantidade de estudantes que foram aprovados em pelo menos um dos cursos e responderam ao questionário; e (vi) a organização e o AVA usados para a pesquisa. Cada um dos possíveis motivos precisa de uma investigação específica para que seja encontrada a causa do problema.

Mesmo não tendo sido possível criar o algoritmo, foi possível observar padrões comportamentais de interação. Descobriu-se que, de forma geral, o comportamento dos estudantes é estável entre cursos quando comparado ao de seus colegas. Isso significa que, apesar do comportamento mudar de um curso para o outro, ele muda de forma relativamente homogênea para todos. Além dessa tendência geral, descobriu-se que aproximadamente 6% dos estudantes variou muito seu comportamento em relação a seus colegas em mais de metade dos indicadores. Ou seja, existe um grupo pequeno de estudantes que, ao mudar de curso, muda completamente seu comportamento, não apenas ajustando-se à tendência do curso. Esse tipo de informação é complementar aos apontamentos da falta de dados sobre a constância do estilo de aprendizagem em estudos longitudinais (Pashler *et al.*, 2009). Para entender de forma mais completa esses padrões comportamentais, é necessário criar novos indicadores, principalmente que tragam informações sobre a regulação, e testá-los, junto com os indicadores já levantados, em outros contextos. Mesmo que as conclusões não sejam generalizáveis, a forma de investigação dos fenômenos proposta nesta dissertação é relevante para profissionais que gerenciam AVAs, pois identificam os padrões comportamentais dos estudantes na plataforma.

Foi possível, no contexto desta pesquisa, identificar agrupamentos de indicadores comportamentais de interação que corroboram com as teorias de estilos de aprendizagem.

Esses agrupamentos não só tiveram boas evidências de validade de construto como também de critério, quando testadas em relação à evasão. Os estilos de aprendizagem estão ligados à evasão, mas seu poder de predição é baixo (Jesus, 2014), uma vez que existem muitas variáveis intervenientes. A predição de 68% da evasão com base nos estilos observados de interação não é muito alta, mas está de acordo com os achados empíricos da área. Esses resultados abrem uma nova linha de investigação em relação à observação de comportamento.

O método proposto para investigar o mesmo construto usando um instrumento autodeclarado e dados observados foi exploratório. Ao analisar os dados, percebeu-se que a forma de usar os dois métodos ao mesmo tempo foi ineficiente. Quando se usa dados comportamentais para inferir um construto autodeclarado são perdidas informações autodeclaradas e observadas. Não é possível observar tudo que é declarado, uma vez que nem tudo se expressa em comportamentos. Por outro lado, muitos fenômenos não declarados são expressos em comportamentos. Sendo assim, a melhor forma de otimizar o uso dos dois métodos seria usá-los de forma complementar, entendendo que eles explicam partes diferentes do fenômeno. Os dois métodos também culminam em investigações práticas diferentes. Por exemplo, a observação de comportamento em AVA apresenta um problema que o questionário não apresenta que é o *cold-start* (Al-Azawei & Badii, 2014). Esse problema é basicamente não se ter o mínimo necessário de informação sobre um indivíduo. Quando se aplica um questionário, as informações necessárias, em geral, já estão no próprio questionário. Por outro lado, definir um padrão de comportamento exige que uma quantidade suficiente de comportamentos seja observada. Além da quantidade mínima, analisar vários cursos simultaneamente apresenta um segundo problema, que é definir uma quantidade equivalente de estímulo ao qual os estudantes são expostos para que seja possível compará-los. Por exemplo, se fosse comparada a quantidade de postagens feitas por um estudante que realizou metade de um curso com um estudante que foi aprovado no curso, haveria uma distorção.

Sendo assim, foi definido como critério que só fossem analisadas respostas de estudantes que foram aprovados no cursos e isso fez com que só fosse possível investigar a evasão dos indivíduos que foram aprovados em pelo menos um curso. Esses indivíduos naturalmente evadem menos e, no caso da amostra do estudo, apenas 5% evadiram. Um ponto positivo do método é que foi possível analisar todos os indivíduos que foram aprovados em pelo menos um curso sem nenhuma intervenção, como, por exemplo, a aplicação de um questionário. No caso dos estilos autodeclarados, foi possível analisar os indivíduos que não foram aprovados em nenhum curso e a proporção de evadidos para essa amostra foi consideravelmente maior (33%). Um ponto negativo do método é que só foi possível analisar os indivíduos que responderam ao questionário.

Por último, a presente Dissertação usou técnicas de análise de dados de *machine learning* que não foram encontradas em nenhum estudo brasileiro da subárea de CO. Essas técnicas foram, nesse caso, superiores às análises tradicionais usadas em CO, uma vez que não foi possível identificar relação entre as variáveis por meio delas. Ainda assim, não foi possível aproveitar todo o potencial dessas poderosas análises, uma vez que a amostra era pequena. A ferramenta WEKA mostrou-se viável para a realização dessas análises.

DISCUSSÃO

Muitas vezes não é possível identificar problemas metodológicos por desconhecimento de formas alternativas de investigar fenômenos ou por falta de pontos de comparação. Os problemas com respostas a escalas são amplamente discutidos na literatura (Schwarz & Oyserman, 2001), no entanto, não é possível quantificar a distorção gerada pelas respostas. Uma vez que isolar um comportamento e usar um delineamento experimental leva à perda de precisão (McGrath, 1981), a alternativa metodológica escolhida para escalas é basicamente compará-las com outras escalas e aplicá-las em diferentes contextos. Apesar de isso conferir alguma segurança quanto à validade de um construto, refere-se apenas a um construto “falado”. Para alguns conceitos, principalmente os mais internos, como, por exemplo, a satisfação com o trabalho, o “dizer” pode estar muito próximo do fenômeno (Mourão *et al.*, 2016). No entanto, para outros fenômenos, principalmente os mais externos, como, por exemplo, a frequência de um comportamento, o “dizer” pode estar muito longe do “fazer”. Nesses casos, a observação do comportamento não só pode trazer uma informação mais fidedigna como também pode trazer uma informação nova, que é a posição relativa do indivíduo em relação à população.

Os comportamentos devem ser investigados diretamente e, ao invés de investiga-los de forma qualitativa ou, no máximo, fazer o percentual e a média deles, como foi descrito no Capítulo 1 que é feito nos artigos brasileiros de CO, eles podem ser investigados com as mesmas análises psicométricas que as escalas. Os estudos de escalas têm uma metodologia robusta para garantir a validade dos instrumentos e essa construção metodológica pode ser usada para a observação de comportamento. A observação de comportamentos usando essa linha metodológica pode ser adotada para vários construtos de CO.

Na presente Dissertação, o que foi encontrado ao analisar estilos de aprendizagem autodeclarados e observados foi que eles são tão diferentes que podem inclusive ser considerados como construtos distintos. Ainda assim, os estilos autodeclarados e os observados trouxeram informações relevantes para a análise dos estudantes em AVA. A relação entre a observação do comportamento usando *big data* e a resposta a escalas é similar à relação entre os métodos quantitativos e qualitativos: trazem informações diferentes sobre os fenômenos, têm vantagens e desvantagens distintas e podem ser usados de forma comparativa ou complementar. A triangulação metodológica e uso de diversos delineamentos em uma mesma investigação são uma tendência em POT (Schein, 2015) e a observação de comportamento usando *big data* é uma metodologia que traz benefícios que não são encontrados em nenhum outro método. Com ele é possível obter dados reais, precisos, generalizáveis e em grande quantidade.

Por isso, há grande potencial para o desenvolvimento de soluções que auxiliem a prática profissional. Por exemplo, a detecção automática de estilos de aprendizagem para um AVA a partir da aprovação em um curso poderia auxiliar professores, tutores e gestores educacionais a potencializarem a aprendizagem e diminuir a evasão. Apesar de não ter sido possível desenvolver essa solução no presente estudo, há indícios de que é possível desenvolvê-la. As análises usando *big data*, apesar de apresentarem uma alta complexidade metodológica, têm um potencial de impacto prático muito grande e devem ser exploradas nos próximos anos por profissionais de POT (Grimm *et al.*, 2017).

Por último, a presente dissertação usou técnicas de análise de dados de *machine learning* que não foram encontradas em nenhum estudo brasileiro da subárea de CO. Considerando que o uso de análises preditivas é uma tendência da subárea de POT (Grimm *et al.*, 2017), os métodos propostos são uma alternativa viável para seguir a tendência. Apesar de não ter sido possível aproveitar todo o potencial desses métodos usando um conjunto maior de

variáveis e uma amostra maior, já foi possível encontrar dois exemplos de seu benefício. Realizando análises comumente usadas em CO para analisar a evasão, no caso, a regressão logística, não foi possível observar nenhum efeito das variáveis independentes na variável dependente, assim como não foi possível criar um modelo de predição melhor do que a distribuição aleatória. Usando técnicas de *machine learning*, foi possível criar um modelo que predissesse a variável critério, para os fatores de estilos autodeclarados e de estilos observados, com precisão significativamente acima do aleatório. Considerando as inúmeras aplicações possíveis dessas técnicas em POT (Chang *et al.*, 2013), é importante que as pesquisas brasileiras as considerem como uma possibilidade.

REFERÊNCIAS

- Abbad, G. S., & Borges-Andrade, J. E. (2004). Aprendizagem humana em organizações de trabalho. *Psicologia, organizações e trabalho no Brasil*, 2, 237-275.
- Abbad, G. S., Mourão, L., Meneses, P. P., Zerbini, T., Borges-Andrade, J. E., & Vilas-Boas, R. (2009). *Medidas de Avaliação em Treinamento, Desenvolvimento e Educação: ferramentas para gestão de pessoas*. Porto Alegre: Artmed Editora.
- Abbad, G., Zerbini, T., & Souza, D. B. L. (2010). Panorama das pesquisas em educação a distância no Brasil. *Estudos de Psicologia* (Natal), 15(3), 291-298.
- Aghaee, M., & Aghaee, R. (2016) Selection of Logistics Personnel by Using and Hybrid Fuzzy DEMATEL and Fuzzy ANP. *International Research Journal of Management Science*, 4(1), 14-22.
- Aguinis, H., Pierce, C. A., Bosco, F. A., & Muslin I. S. (2009). First decade of organizational research methods: trends in design, measurement, and data-analysis topics. *Organizational Research Methods* 12(1), 69-112. <http://doi.org/b2jd4d>.
- Al-Azawei, A., & Badii, A. (2014). State of the art of learning styles-based adaptive educational hypermedia systems (LS-BAEHSs). *International Journal of Computer Science & Information Technology*, 6(3), 1-19. <http://doi.org/cz5m>.
- Alkhuraiji, S. M. (2016). *Dynamic adaptive e-learning mechanism based on learning styles*. Tese de Doutorado, Universidade de Manchester, Manchester, UK.
- Allinson, C. W., & Hayes, J. (1996). The cognitive style index: A measure of intuition-analysis for organizational research. *Journal of Management studies*, 33(1), 119-135.

<http://doi.org/bn39c7>.

Amir, E. S., Sumadyo, M., Sensuse, D. I., Sucahyo, Y. G., & Santoso, H. B. (2016, Outubro).

Automatic detection of learning styles in learning management system by using literature-based method and support vector machine. *Proceedings of the Advanced Computer Science and Information Systems (ICACSIS) International Conference*.

Malang, Indonésia. <http://doi.org/cz5n>.

Angrave, D., Charlwood, A., Kirkpatrick, I., Lawrence, M., & Stuart, M. (2016). HR and

analytics: why HR is set to fail the big data challenge. *Human Resource Management Journal*, 26(1), 1-11. <http://doi.org/c3cz>.

Apter, M. J., Mallows, R., & Williams, S. (1998). The development of the motivational style

profile. *Personality and Individual Differences*, 24(1), 7-18. <http://doi.org/bd9mmh>.

Arslan, A. K. (2018). An exploratory model of learning styles based on agent

learning. *Advances in Higher Education*, 2(2). <http://doi.org/cz5p>.

Azar, A., Sebt, M. V., Ahmadi, P., & Rajaeian, A. (2013). A model for personnel selection

with a data mining approach: A case study in a commercial bank. *SA Journal of Human Resource Management*, 11(1), 1-10. <http://doi.org/cxm8>.

Azevedo, R., Harley, J., Trevors, G., Duffy, M., Feyzi-Behnagh, R., Bouchet, F., & Landis, R.

(2013). Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In *International handbook of metacognition and learning technologies* (pp. 427-449). New York, NY: Springer. <http://doi.org/cxm9>.

Bajraktarevic, N., Hall, W., & Fullick, P. (2003, Agosto). Incorporating learning styles in

- hypermedia environment: Empirical evaluation. *Proceedings of the workshop on adaptive hypermedia and adaptive web-based systems*. Nottingham, UK.
- Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In *Learning analytics* (pp. 61-75). New York: Springer. <http://doi.org/cz5q>.
- Bastos, A. V. B., Maia, L. G., Rodrigues, A. C. A., Macambira, M. O. & Borges-Andrade, J. E. (2014). Vínculos dos indivíduos com a organização: Análise da produção científica brasileira 2000-2010. *Psicologia: Teoria e Pesquisa*, 30 (2), 153-162. <http://doi.org/cxmw>.
- Begum, R. M., & David, K. (2017). Discovering Student Learning Style using Min Max Cascade Neural Network. *Indian Journal of Science and Technology*, 10(25). <http://doi.org/cz5r>.
- Bell, B. S., Tannenbaum, S. I., Ford, J. K., Noe, R. A., & Kraiger, K. (2017). 100 years of training and development research: What we know and where we should go. *Journal of Applied Psychology*, 102(3), 305. <http://doi.org/f9xt5j>.
- Bendassolli, P. F., Borges-Andrade, J. E., & Malvezzi, S. (2010). Paradigmas, eixos temáticos e tensões na PTO no Brasil. *Estudos de Psicologia (Natal)*, 15 (3), 281–289. <http://doi.org/fm9fg3>.
- Berings, M. G., Poell, R. F., & Simons, P. R. J. (2005). Conceptualizing on-the-job learning styles. *Human Resource Development Review*, 4(4), 373-400. <http://doi.org/bwx7n2>.
- Bernard, J., Chang, T. W., Popescu, E., & Graf, S. (2016). Improving Learning Style Identification by Considering Different Weights of Behavior Patterns Using Particle Swarm Optimization. In *State-of-the-Art and Future Directions of Smart Learning* (pp.

- 39-49). Singapore: Springer. <http://doi.org/cz5s>.
- Bernard, J., Chang, T. W., Popescu, E., & Graf, S. (2015, Junho). Using artificial neural networks to identify learning styles. In *International Conference on Artificial Intelligence in Education* (pp. 541-544). Cham, Switzerland: Springer. <http://doi.org/cz5t>.
- Biggs, J. B. (1987). *Student Approaches to Learning and Studying*. Hawthorn, Victoria: Australian Council for Educational Research.
- Borges-Andrade, J. E., Abbad, G. S., & Mourão, L. (2006). *Treinamento, desenvolvimento e educação em organizações e trabalho: fundamentos para a gestão de pessoas*. Porto Alegre: Artmed.
- Borges-Andrade, J. E., Renteria-Pérez, E., & Toro, J. P. (2018). Organizational/Work Psychology in Latin America. In R., Ardila. *Psychology in Latin America* (pp. 105-158). Cham, Switzerland: Springer. <http://doi.org/cxnb>.
- Brusilovsky, P. (1994, Agosto). Student model centered architecture for intelligent learning environments. In *Proceedings of Fourth International Conference on User Modeling* (pp. 31-36). Hyannis, USA: Kluwer Academic Publishers .
- Cairns, L., & Malloch, M. (2011). Theories of work, place and learning: New directions. *The Sage handbook of workplace learning*, (pp. 3 – 16). Londres: Sage. <http://doi.org/cz5v>.
- Cantal, C., Borges-Andrade, J. E & Porto, J. B. (2015). Cooperação, comportamentos proativos, ou simplesmente cidadania organizacional? Uma revisão da produção nacional na área. *Revista Psicologia: Organizações e Trabalho*, 15 (3), 286-297. <http://dx.doi.org/http://doi.org/cxmx>.

- Carver, C. A., Howard, R. A., & Lane, W. D. (1999). Addressing different learning styles through course hypermedia. *IEEE Transactions on Education*, 42(1), 33-38.
- Chang, R. M., Kauffman, R. J., & Kwon, Y. (2013). Understanding the paradigm shift to computational social science in the presence of big data. *Decision Support Systems*, 63, 67-80. <http://doi.org/f584j6>.
- Coelho, F. A. Jr., , Borges-Andrade & J. E. (2011). Efeitos de variáveis individuais e contextuais sobre desempenho individual no trabalho. *Estudos de Psicologia*, 16(2), 111-120. <http://doi.org/c3c2>.
- Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). *Learning styles and pedagogy in post-16 learning. A systematic and critical review*. London: Learning and Skills Research Centre.
- Damásio, B. F. (2012). Uso da análise fatorial exploratória em psicologia. *Avaliação Psicológica: Interamerican Journal of Psychological Assessment*, 11(2), 213-228.
- Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning*. Harvard Business Press.
- De Bello, T. C. (1990). Comparison of eleven major learning styles models: Variables, appropriate populations, validity of instrumentation, and the research behind them. *Reading, Writing, and Learning Disabilities*, 6(3), 203-222.
- Deloitte (2017). *Global human capital trends: Rewriting the rules for the digital age*. Retrieved from https://www.lbg-canada.ca/wp-content/uploads/2017/10/Michelle-Ryder_Global-HC-Trends-2017-Insideout-Nov-2-2017.pdf.
- Dorça, F. A. (2012). Uma abordagem estocástica baseada em aprendizagem por reforço para

modelagem automática e dinâmica de estilos de aprendizagem de estudantes em sistemas adaptativos e inteligentes para educação a distância. Tese de Doutorado, Universidade Federal de Uberlândia, Uberlândia, MG.

Dunn, R. S., Dunn, K. J., & Price, G. E. (1981). *Learning style inventory*. New York, Lawrence, KS: Price Systems.

El Aissaoui, O., El Madani, Y. E. A., Oughdir, L., & El Alloui, Y. (2018) A fuzzy classification approach for learning style prediction based on web mining technique in e-learning environments. *Education and Information Technologies* (pp. 1 – 17). <http://doi.org/cz5w>.

Englert, P. (2012). Locally Weighted Learning. *Seminar Class on Autonomous Learning Systems*.

Entwistle, N. J. (1981). *Styles of learning and teaching: An integrated outline of educational psychology*. Clichester: Wiley.

Entwistle N (1990). Teaching and the quality of learning in higher education. Em N Entwistle (ed.) *Handbook of educational ideas and practices*. London: Routledge.

Entwistle, N. J., & Peterson, E. R. (2004). Conceptions of learning and knowledge in higher education: Relationships with study behaviour and influences of learning environments. *International journal of educational research*, 41(6), 407-428. <http://doi.org/b4ng2n>.

Entwistle, N., McCune, V., & Tait, H. (2013). Approaches and Study Skills Inventory for Students (ASSIST) (incorporating the Revised Approaches to Studying Inventory-RASI). *Recuperado de*

https://www.researchgate.net/publication/260291730_Approaches_and_Study_Skills_Inventory_for_Students_ASSIST_incorporating_the_Revised_Approaches_to_Studying_Inventory_-_RASI.

Fasihuddin, H., Skinner, G., & Athauda, R. (2017). Towards adaptive open learning environments: Evaluating the precision of identifying learning styles by tracking learners' behaviours. *Education and Information Technologies*, 22(3), 807-825. <http://doi.org/cz5x>.

Fatahi, S., Shabanali-Fami, F., & Moradi, H. (2017). An empirical study of using sequential behavior pattern mining approach to predict learning styles. *Education and Information Technologies*, 1-19. <http://doi.org/cz5z>.

Felder, R. M., & Silverman, L. K. (1988). Learning and teaching styles in engineering education. *Engineering education*, 78(7), 674-681.

Felder, R. M. (1996). Matters of style. *ASEE prism*, 6(4), 18-23.

Felder, R. M., & Spurlin, J. (2004). Applications, reliability and validity of the index of learning styles. *International journal of engineering education*, 21(1), 103-112.

Feldman, J., Monteserin, A., & Amandi, A. (2015). Automatic detection of learning styles: state of the art. *Artificial Intelligence Review*, 44(2), 157-186. <http://doi.org/f7kn99>.

Ferreira, L., Gazzola, M., Ferrari, D., Zupelari, M., Paiva, P., & Rodrigues Jr, J. (2017, Outubro). Métodos de classificação aplicados à detecção automática de Estilos de Aprendizagem em um ambiente real de ensino. *Anais do Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)* (p. 1517). Recife, PE, Brasil.

- Fonseca, A. M. O., Porto, J. B. & Borges-Andrade, J. E (2015). Liderança: Um retrato da produção científica brasileira. *Revista de Administração Contemporânea*, 19 (3), 290-310. <http://doi.org/cxmz>.
- Gagne, R. M., & Briggs, L. J. (1974). *Principles of instructional design*. Holt, Rinehart & Winston.
- Gardner, H. (1983). *Frames of mind: the Idea of multiple intelligence*. New York, NY.
- Graf, S. (2007). *Adaptivity in learning management systems focussing on learning styles*. Tese de Doutorado, Faculty of Informatics, Vienna University of Technology, Viena, Áustria.
- Gregorc, A. F. (1979). Learning/teaching styles: Their nature and effects. *Student learning styles: Diagnosing and prescribing programs*, 19-26.
- Grimm, K. J., Jacobucci, R., & McArdle, J. J. (2017, Janeiro). *Big data methods and psychological science*. Recuperado em 4 de dezembro, 2018, de: <https://www.apa.org/science/about/psa/2017/01/big-data-methods.aspx>.
- Guzzo, R. A., Fink, A. A., King, E., Tonidandel, S., & Landis, R. S. (2015). Big data recommendations for industrial–organizational psychology. *Industrial and Organizational Psychology*, 8(4), 491-508. <http://doi.org/cxnc>.
- Hart, C. (2012). Factors associated with student persistence in an online program of study: A review of the literature. *Journal of Interactive Online Learning*, 11(1).
- Hiltz, S. R. (1994). *The virtual classroom: Learning without limits via computer networks*. Intellect Books.
- Honey, P., & Mumford, A. (1992). *The manual of learning styles*.

- Hutcheson, G. D. & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. London: Sage Publications
- Jackson, C. (2002). *Manual of the learning styles profiler*. Sydney: Cymeon Research.
- Jesus, S. C. B. P. de. (2014). *Estudo Multinível de Antecedentes do Impacto no Trabalho e da Persistência/Evasão em Treinamentos a Distância*. Tese de Doutorado, Psicologia Social, do Trabalho e das Organizações (PSTO), Universidade de Brasília, Brasília, Brasil.
- Jaiswal, A. K., Singh, N., & Ahuja, N. J. (2017). Learning styles based adaptive intelligent tutoring systems: Document analysis of articles published between 2001 and 2016. *International Journal of Cognitive Research in Science, Engineering and Education: (IJCRSEE)*, 5(2), 83-98. <http://doi.org/cz52>.
- Kagan, J., & Kogan, N. (1970). Individual variation in cognitive processes. *Carmichael's manual of child psychology* (3a ed.). New York: Wiley.
- Kalhor, A. A., Rajper, S., & Mallah, G. A. (2016). Detection of E-Learners' Learning Styles: An Automatic Approach using Decision Tree. *International Journal of Computer Science and Information Security (IJCSIS)*, 14(8).
- Karagiannis, I., & Satratzemi, M. (2018). An adaptive mechanism for Moodle based on automatic detection of learning styles. *Education and Information Technologies*, 23(3), 1331-1357. <http://doi.org/cz53>.
- Kauffman, H. (2015). A review of predictive factors of student success in and satisfaction with online learning. *Research in Learning Technology*, 23. <http://doi.org/cz54>.
- Keefe, J. W. (ed.) 1979. *Student Learning Styles: Diagnosing and Prescribing Programs*.

Reston, VA: National Association of Secondary School Principals.

Kirimi, J. M., & Moturi, C. A. (2016). Application of Data Mining Classification in Employee Performance Prediction. *International Journal of Computer Applications (0975-8887)*, 146(7). <http://doi.org/cz55>.

Kolb, D. A. (1976). *Learning style inventory technical manual*. Boston, MA: McBer.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice-Hall, Inc., Englewood Cliffs, N.J.

Kraiger, K. (2008). Third-generation instructional models: More about guiding development and design than selecting training methods. *Industrial and Organizational Psychology*, 1(4), 501-507. <http://doi.org/fv3g9m>.

Leydesdorff, L., Zhou, P., & Bornmann, L. (2013). How can journal impact factors be normalized across fields of science? An assessment in terms of percentile ranks and fractional counts. *Journal of the American Society for Information Science and Technology*, 64(1), 96-107. <http://doi.org/gffnx5>.

Li, L. X., & Abdul Rahman, S. S. (2018). Students' learning style detection using tree augmented naive Bayes. *Royal Society open science*, 5(7), 172108. <http://doi.org/cz56>.

Lichtenstein, S., & Slovic, P. (Eds.). (2006). *The construction of preference*. Cambridge University Press.

Liu, J., & Moskvina, A. (2016). Hierarchies, ties and power in organizational networks: model and analysis. *Social Network Analysis and Mining*, 6(1), 106. <http://doi.org/gdj8c8>.

- Lugo, G. S., Rodríguez, L. F., López, R. I. G., Macías-Estrada, A., & Echeverría, M. R. (2015). Behavioral Patterns for Automatic Detection of Learning Styles in Learning Management Systems: a Case Study. *Research in Computing Science*, 106, 69-77.
- McGrath, J. E. (1981). Dilemmatics: The study of research choices and dilemmas. *American Behavioral Scientist*, 25(2), 179-210. <http://doi.org/fkf8gr>.
- Moraes, R. B. N. (2016). *Estilos de aprendizagem em ações educacionais ofertadas a distância: evidências de validade, validade convergente e análise conceitual*. Dissertação de Mestrado, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto (FFCLRP), Universidade de São Paulo, Ribeirão Preto, Brasil.
- Moraes, R. B. N., & Zerbini, T. (2018). Estilos de Aprendizagem em EaD: Construção e evidências de validade de instrumento. *Estudos e Pesquisas em Psicologia*, 18(1).
- Mortenson, M. J., Doherty, N. F., & Robinson, S. (2015). Operational research from Taylorism to Terabytes: A research agenda for the analytics age. *European Journal of Operational Research*, 241(3), 583-595. <http://doi.org/f8mk5z>.
- Mourão, L., Bastos, A. V. B., & Oliveira, R. P. D. (2016). O dizer eo fazer nas pesquisas em POT. *Revista Psicologia Organizações e Trabalho*, 16(4), 333-339. <http://doi.org/cxm2>.
- Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator* (Vol. 3). Palo Alto, CA: Consulting Psychologists Press.
- Nascimento, T. T., Borges-Andrade, J. E & Porto, J. B. (2016). Produção científica brasileira sobre comportamento organizacional no terceiro setor. *Estudos de Psicologia (Campinas)*, 33 (2), 367-374. <http://doi.org/cxnd>.

- Noe, R. A., Clarke, A. D., & Klein, H. J. (2014). Learning in the twenty-first-century workplace. *Annu. Rev. Organ. Psychol. Organ. Behav.*, *1*(1), 245-275.
<http://doi.org/gckf46>.
- de Oliveira, E. S., Sales, G. L., de Sousa Pereira, P., & do Nascimento Moreira, R. (2018, Julho). Identificação Automática de Estilos de Aprendizagem: Uma Revisão Sistemática da Literatura. In *Workshop sobre Educação em Computação (WEI_CSBC) 26(1/2018)*.
- Paoletti, A. L., Martinez-Gil, J., & Schewe, K. D. (2015, September). Extending knowledge-based profile matching in the human resources domain. In *International Conference on Database and Expert Systems Applications* (pp. 21-35). Springer, Cham.
<http://doi.org/cxm3>.
- Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society*, *17*(4).
- Papanikolaou, K. A., Grigoriadou, M., Kornilakis, H., & Magoulas, G. D. (2003). Personalizing the Interaction in a Web-based Educational Hypermedia System: the case of INSPIRE. *User modeling and user-adapted interaction*, *13*(3), 213-267.
<http://doi.org/dbx9jm>.
- Paredes, P., & Rodriguez, P. (2004). A mixed approach to modelling learning styles in adaptive educational hypermedia. *Advanced Technology for Learning*, *1*(4), 210-215.
<http://doi.org/b9kfb9>.
- Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2009). Learning styles: Concepts and evidence. *Psychological science in the public interest*, *9*(3), 105-119.
<http://doi.org/cktcsf>.

- Pasquali, L. (2009). Psychometrics. *Revista da Escola de Enfermagem da USP*, 43(SPE), 992-999.
- Pasquali, L. (2010). Testes referentes a construto: Teoria e modelo de construção. In L. Pasquali (Ed.), *Instrumentação Psicológica: Fundamentos e Prática* (pp. 165-188) Porto Alegre: Artmed.
- Petchboonmee, P., Phonak, D., & Tiantong, M. (2015). A comparative data mining technique for david kolb's experiential learning style classification. *International Journal of Information and Education Technology*, 5(9), 672. <http://doi.org/cz57>.
- Popescu, E. (2008). *Dynamic adaptive hypermedia systems for e-learning*. Tese de Doutorado, Université de Technologie de Compiègne, Compiègne, France.
- Popescu, E. (2009). Diagnosing students' learning style in an educational hypermedia system. In *Cognitive and emotional processes in Web-based education: Integrating human factors and personalization* (pp. 187-208). IGI Global. <http://doi.org/bpxfbx>.
- Porter, L. W., & Schneider, B. (2014). What Was , What Is , and What May Be in OP / OB. *Annual Review of Organizational Psychology and Organizational Behavior*, 1(1), 1–21. <http://doi.org/cxm4>.
- Puente-Palacios, K., & Peixoto, A. D. L. A. (2015). *Ferramentas de Diagnóstico para Organizações e Trabalho: Um Olhar a partir da Psicologia*. Porto Alegre, Brasil: Artmed Editora.
- Rajper, S., Shaikh, N. A., Shaikh, Z. A., & Mallah, G. A. (2016). Automatic detection of learning styles on learning management systems using data mining technique. *Indian Journal of Science and Technology*, 9(15). <http://doi.org/cz58>.

- Resende, S. & Neiva, E. R. (2018). Avaliação dos instrumentos de medida de afeto e bem-estar no trabalho publicados em periódicos de psicologia e administração de 1996 a 2015. *Trabalho (En) Cena*, 3 (2), 80-94. <http://doi.org/cxm5>.
- Rodrigues, A. C. de A., & Carvalho-Freitas, M. N. de. (2016). Theoretical fragmentation : Origins and repercussions in Work and Organizational. *Revista Psicologia: Organizações e Trabalho*, 16(4), 310–315. <http://doi.org/cxm6>.
- Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(6), 601-618. <http://doi.org/ctnbcs>.
- Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12-27. <http://doi.org/cz59>.
- Sahid, D. S. S., Nugroho, L. E., & Santosa, P. I. (2017). Integrated Stochastic and Literate Based Driven Approaches in Learning Style Identification for Personalized E-Learning Purpose. *International Journal on Advanced Science, Engineering and Information Technology*, 7(5), 1708-1715. <http://doi.org/cz6b>.
- Salles, T. (2007). *Estilos de aprendizagem no trabalho: análise e construção de medidas*. Dissertação de Mestrado. Universidade de Brasília. Brasília, DF.
- Sampaio, N. S. P, Borges-Andrade, J. E., & Bonatti, C. L. (2018). Aprendizagem no trabalho: pesquisa nas organizações públicas e privadas brasileiras. *Psicología Desde el Caribe*, 35 (n.esp.), prelo.
- Sena, E. B., Vivas, A., Assis, L., & Pitanguí, C. (2016). Uma abordagem computacional para detecção automática de estilos de aprendizagem utilizando modelos ocultos de markov.

In *SIED: EnPED-Simpósio Internacional de Educação a Distância e Encontro de Pesquisadores em Educação a Distância*. São Carlos, SP, Brasil.

Schein, E. H. (2015). Organizational Psychology Then and Now : Some Observations. *Annual Review of Organizational Psychology and Organizational Behavior*, 2(1), 1–19.

<http://doi.org/cxm7>.

Schwarz, N., & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. *American Journal of Evaluation*, 22(2), 127-160. <http://doi.org/drw7jk>.

Sheeba, T., & Krishnan, R. (2018, Julho). Prediction of student learning style using modified decision tree algorithm in e-learning system. In *Proceedings of the 2018 International Conference on Data Science and Information Technology* (pp. 85-90). Singapore: ACM. <http://doi.org/cz6c>.

Siemens, G., & Gasevic, D. (2012). Guest editorial-Learning and knowledge analytics. *Educational Technology & Society*, 15(3), 1-2.

Siemens, G., & Baker, R. S. (2012, Abril). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252-254). Vancouver, Canada: ACM.

Siqueira, M. M. M. (2002). Medidas do comportamento organizacional. *Estudos de Psicologia*, 7 (Número Especial), 11-18.

Siqueira, M. M. M. (2009). *Medidas do comportamento organizacional: ferramentas de diagnóstico e de gestão*. Porto Alegre, Brasil: Artmed Editora.

- Siqueira, M. M. M. (2014). *Novas medidas do comportamento organizacional: ferramentas de diagnóstico e de gestão*. Porto Alegre, Brasil: Artmed Editora.
- Stash, N., Cristea, A., & De Bra, P. (2006, Outubro). Adaptation to learning styles in e-learning: Approach evaluation. In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 284-291). Chesapeake, VA: AACE.
- Stathacopoulou, R., Magoulas, G. D., Grigoriadou, M., & Samarakou, M. (2005). Neuro-fuzzy knowledge processing in intelligent learning environments for improved student diagnosis. *Information Sciences*, 170(2-4), 273-307. <http://doi.org/ck8qss>.
- Sternberg, R. J. (1999). *Thinking styles*. Cambridge university press.
- Sternberg, R. J., Zhang, L. F., & Rayner, S. (Eds.). (2011). *Handbook of intellectual styles: Preferences in cognition, learning, and thinking*. Springer Publishing Company.
- Sweta, S., & Lal, K. (2015, Dezembro). Web Usages Mining in Automatic Detection of Learning Style in Personalized e-Learning System. In *Proceedings of the Fifth International Conference on Fuzzy and Neuro Computing (FANCCO-2015)* (pp. 353-363). Springer, Cham. <http://doi.org/cz6f>.
- Sweta, S., & Lal, K. (2016). Learner Model for Automatic Detection of Learning Style Using FCM in Adaptive E-Learning System. *IOSR J.(IOSR J. Comput. Eng.)*, 18(2), 18-24. <http://doi.org/dswvvsq>.
- Sweta, S., & Lal, K. (2017). Personalized adaptive learner model in e-learning system using FCM and fuzzy inference system. *International Journal of Fuzzy Systems*, 19(4), 1249-1260. <http://doi.org/cz6g>.

- Tait, H., Entwistle, N.J., & McCune, V. (1998). *Approaches and study skills inventory for students (ASSIST)*. Recuperado de www.etl.tla.ed.uk/questionnaires/assist.pdf
- Tseng, J. C., Chu, H. C., Hwang, G. J., & Tsai, C. C. (2008). Development of an adaptive learning system with two sources of personalization information. *Computers & Education, 51*(2), 776-786. <http://doi.org/c8z3cg>.
- Vermunt, J. D. (1994). Inventory of learning styles in higher education. *Maastricht: Maastricht University*.
- Vermunt, J. D. (1996). Metacognitive, cognitive and affective aspects of learning styles and strategies: a phenomenographic analysis. *Higher Education, 31*, 25-50.
- Vermunt, J. D. (1998). The regulation of constructive learning processes. *British Journal of Educational Psychology, 68*, 149-171.
- Wilson, M. L. (2012). Learning styles, instructional strategies, and the question of matching: A literature review. *International Journal of Education, 4*(3), 67-87. <http://doi.org/dxxcth>.
- Witten, I. H., Frank, E., Trigg, L. E., Hall, M. A., Holmes, G., & Cunningham, S. J. (1999). *Weka: Practical machine learning tools and techniques with Java implementations*.
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.
- Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences, 112*(4), 1036-1040. <http://doi.org/gc5p3m>.

APÊNDICE A - Itens classificados por juízes

Item	Categoria	Subcategoria 1	Subcategoria 2	Observável
Foco em ideias.	Processamento cognitivo 63%	Processamento cognitivo 63%		<i>não</i>
Realizo apenas as atividades que são pontuadas.	Regulação da aprendizagem 88%	Estudo organizado 38%	Potencializar efetividade 38%	<i>sim</i>
Pratico o que aprendi.	Regulação da aprendizagem 100%	Potencializar efetividade 88%		<i>não</i>
Leio todas as respostas a postagens que faço	Interação 63%	Aprendizagem vicariante 50%		<i>sim</i>
Divido os materiais de curso de forma a criar um ritmo de estudo que siga até o final do curso	Regulação da aprendizagem 100%	Estudo organizado 63%		<i>sim</i>
Busco materiais além dos que são passados no curso	Regulação da aprendizagem 100%	Potencializar efetividade 100%		<i>não</i>
Reflico sobre o conteúdo para melhor compreendê-lo.	Processamento cognitivo 75%	Processamento cognitivo 75%		<i>não</i>
Interajo com colegas durante o curso	Interação 88%	Interação ativa 63%		<i>sim</i>
Uso minha intuição.	Processamento cognitivo 100%	Processamento cognitivo 100%		<i>não</i>
Acesso os materiais do curso várias vezes	Regulação da aprendizagem 100%	Potencializar efetividade 75%		<i>sim</i>
Acesso os materiais do curso na ordem que são apresentados	Regulação da aprendizagem 100%	Estudo organizado 63%		<i>sim</i>
Verifico minhas notas assim que elas são lançadas.	Regulação da aprendizagem 88%	Potencializar efetividade 63%		<i>sim</i>
Interajo com o tutor durante o curso	Interação 100%	Interação ativa 88%		<i>sim</i>
Quando tenho uma dúvida, pergunto a colegas.	Interação 100%	Interação ativa 100%		<i>sim</i>
Leio todos os comentários sobre as tarefas e avaliações que entrego	Interação 63%	Aprendizagem vicariante 63%		<i>sim</i>
Quando tenho uma dúvida, posto-a no curso.	Interação 88%	Interação ativa 63%		<i>sim</i>

Revejo os materiais do curso até fixar o conteúdo	Regulação da aprendizagem	100%	Potencializar efetividade	63%		não
Acesso os materiais do curso na ordem que considero melhor	Regulação da aprendizagem	100%	Estudo organizado	63%		sim
Busco fatos e dados.	Regulação da aprendizagem	88%	Potencializar efetividade	38%	Estudo organizado 38%	não
Prefiro decidir sozinho como estudar	Processamento cognitivo	63%	Processamento cognitivo	63%		não
Entrego todas as tarefas que têm prazo de entrega com antecedência	Regulação da aprendizagem	100%	Gerenciamento de tempo	88%		sim
Estabeleço minhas próprias metas.	Regulação da aprendizagem	75%	Potencializar efetividade	50%		não
Acesso os materiais adicionais do curso	Regulação da aprendizagem	100%	Potencializar efetividade	50%	Estudo organizado 50%	sim
Acesso o curso sempre no mesmo horário	Regulação da aprendizagem	100%	Gerenciamento de tempo	88%		sim
Utilizo uma técnica específica para realização das atividades.	Regulação da aprendizagem	100%	Estudo organizado	50%	Potencializar efetividade 50%	não
Interajo com colegas e com o tutor sobre assuntos que não estão diretamente ligados ao curso.	Interação	100%	Interação ativa	100%		não
Postagens do tutor e de meus colegas são importantes para meu estudo	Interação	88%	Aprendizagem vicariante	63%		não
A interação com o tutor é importante para que eu aprenda o conteúdo do curso	Interação	75%	Interação ativa	50%		não
Realizo as avaliações que têm prazo em cima da hora	Regulação da aprendizagem	100%	Gerenciamento de tempo	100%		sim
Quando tenho uma dúvida, pergunto ao tutor.	Interação	100%	Interação ativa	100%		sim
Realizo as avaliações com calma para ter o melhor resultado possível	Regulação da aprendizagem	88%	Potencializar efetividade	50%		sim
Uso uma abordagem lógica.	Processamento cognitivo	63%	Processamento cognitivo	63%		não
Realizo um curso de cada vez	Regulação da aprendizagem	100%	Estudo organizado	63%		sim

Realizo vários cursos ao mesmo tempo de áreas correlatas.	Regulação da aprendizagem	100%	Estudo organizado	38%	Potencializar efetividade	25%	não
Faço esquemas e resumos enquanto estudo os conteúdos.	Regulação da aprendizagem	100%	Estudo organizado	63%			não
Crio planos de ação.	Regulação da aprendizagem	88%	Estudo organizado	63%			não
Tento aprender sozinho mesmo quando acho o conteúdo difícil.	Regulação da aprendizagem	63%	Potencializar efetividade	38%	Estudo organizado	25%	não
Prefiro fazer um planejamento de estudos para realizar o curso com sucesso.	Regulação da aprendizagem	100%	Estudo organizado	38%	Potencializar efetividade	25%	não
Entrego as tarefas com prazo de entrega em cima da hora	Regulação da aprendizagem	100%	Gerenciamento de tempo	100%			sim
Estabeleço metas de desempenho no curso.	Regulação da aprendizagem	88%	Potencializar efetividade	63%			não
Olho postagens de colegas e do tutor para ajudar em meu estudo	Interação	88%	Aprendizagem vicariante	88%			sim
Realizo todas as atividades mesmo as que não são pontuadas.	Regulação da aprendizagem	100%	Potencializar efetividade	50%			sim
Ao receber uma nota, vejo quais questões acertei, quais errei e onde errei.	Regulação da aprendizagem	100%	Potencializar efetividade	100%			sim
Aprendo a partir de insights.	Processamento cognitivo	100%	Processamento cognitivo	100%			não
Realizo as avaliações que têm prazo com antecedência	Regulação da aprendizagem	100%	Gerenciamento de tempo	100%			sim
Logo antes das avaliações, reviso os materiais de estudo	Regulação da aprendizagem	100%	Gerenciamento de tempo	50%			sim
Primeiro compreendo a estrutura geral do conteúdo para depois estudá-lo em detalhes.	Regulação da aprendizagem	88%	Estudo organizado	75%			não
A interação com colegas é importante para que eu aprenda o conteúdo do curso	Interação	88%	Interação ativa	63%			não
Acesso todos os materiais de estudo antes de realizar as avaliações	Regulação da aprendizagem	100%	Estudo organizado	50%			sim
Penso em diferentes possibilidades.	Processamento cognitivo	88%	Processamento cognitivo	88%			não

Realizo vários cursos ao mesmo tempo de áreas diferentes.	Regulação da aprendizagem 88%	Estudo organizado 50%		<i>não</i>
Presto atenção aos detalhes antes de chegar a uma conclusão.	Processamento cognitivo 50%	Processamento cognitivo 50%	Estudo organizado 38%	<i>não</i>
Quando tenho uma dúvida, procuro a resposta em postagens e comentários de colegas ou do tutor.	Interação 75%	Aprendizagem vicariante 75%		<i>sim</i>

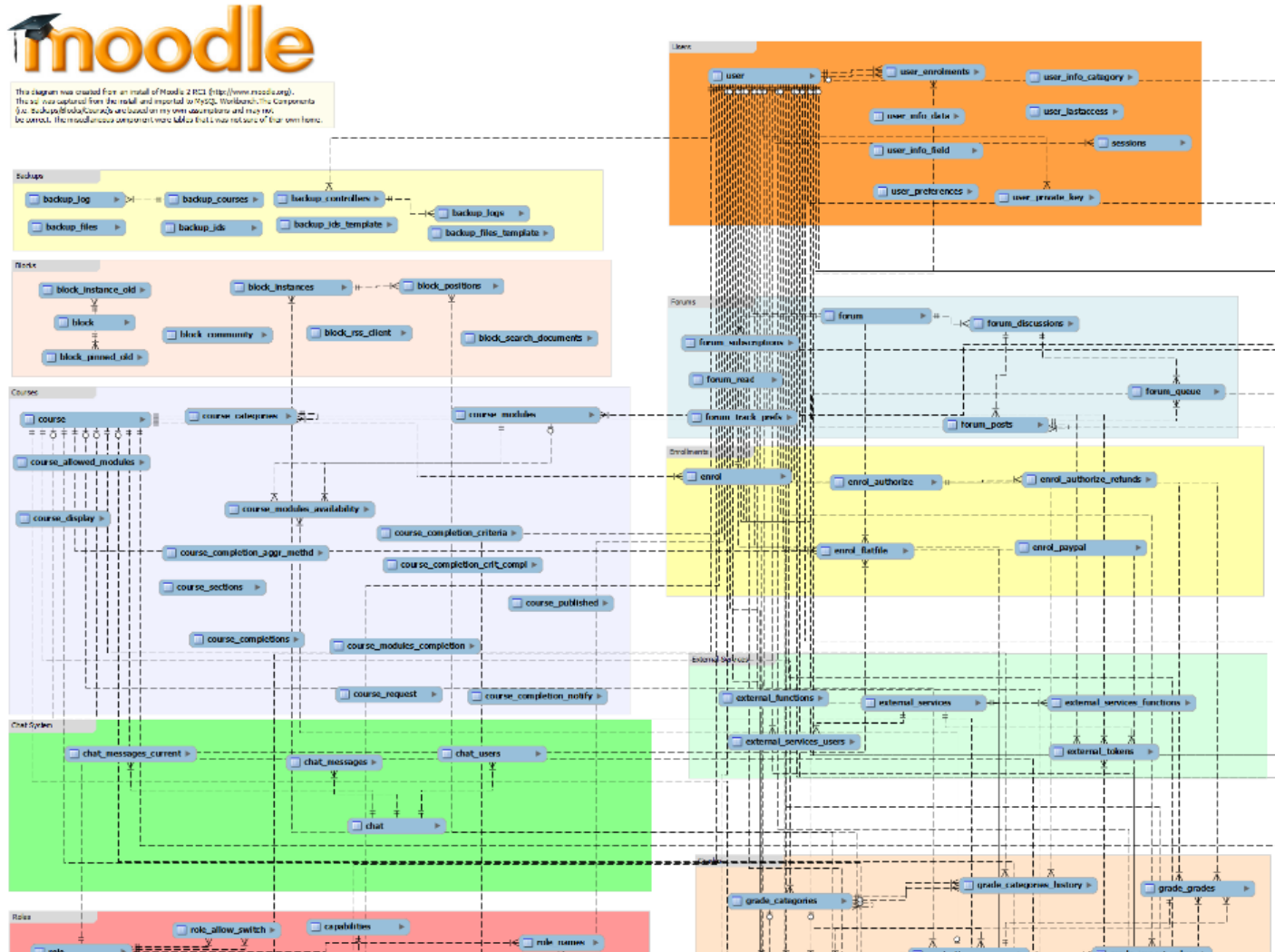
**APÊNDICE B – Revisão de artigos de detecção automática de estilos de aprendizagem
de 2015 a 2018**

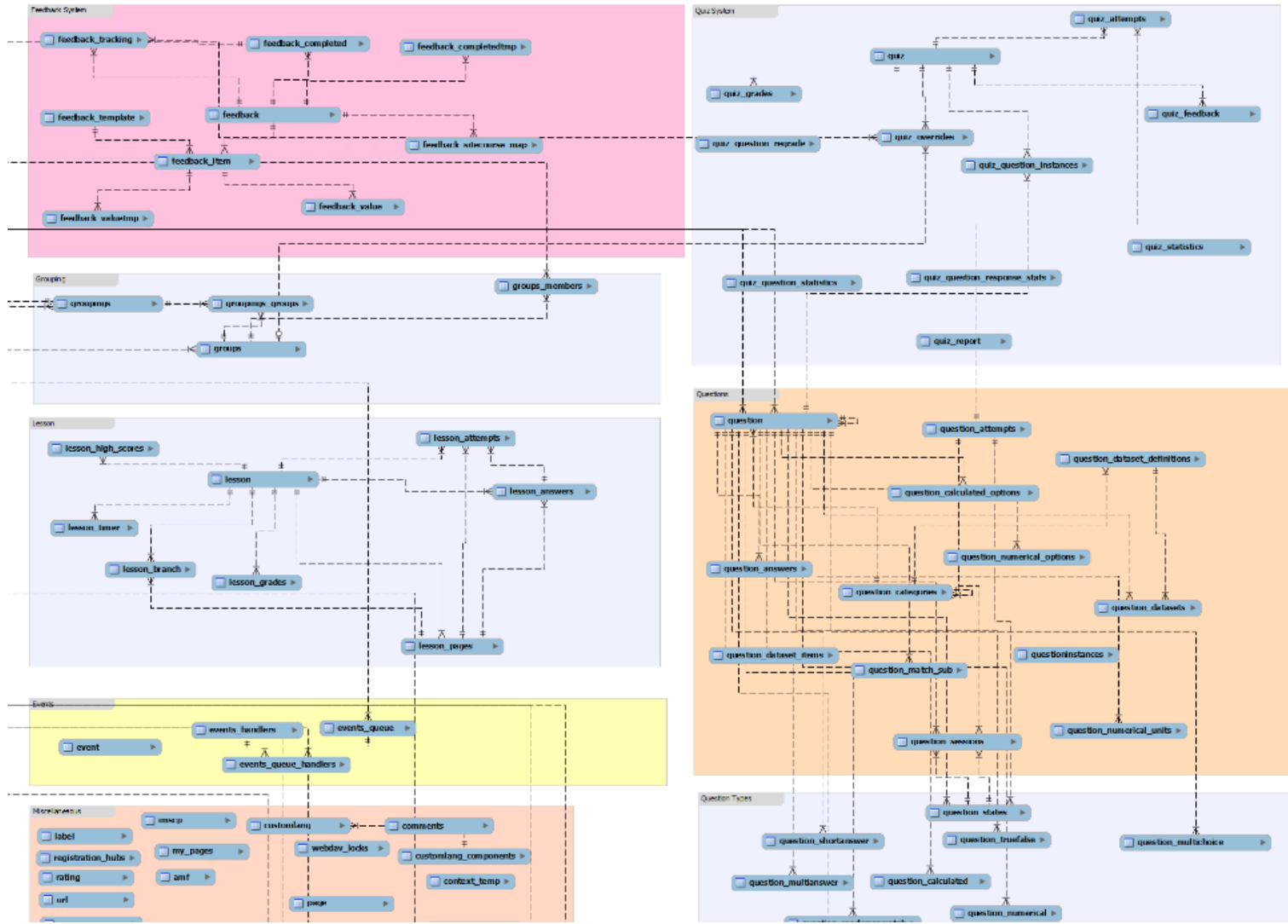
Autores	Modelo de Estilo de Aprendizagem	Abordagem	Método de análise de dados	Quantidade de cursos	Quantidade de estudantes
Alkhuraiji (2016)	Felder-Silverman	Implícita, baseada em dados e dinâmica	Algoritmo de Similaridade	-	-
Amir et al. (2016)		Implícita baseada em literatura e estática	Máquina de Vetores de Suporte	1	200
Arslan (2018)	VARK	Implícita, mista e dinâmica	-	-	-
Begum e David (2017)	Kolb	Implícita, baseada em dados e estática	Redes Neurais Fuzzy	3	180
Bernard, Chang, Popescu e Graf (2015)	Felder-Silverman	Implícita, baseada em literatura e estática	Redes Neurais	1	75
Bernard, Chang, Popescu e Graf (2016)	Felder-Silverman	Implícita, baseada em literatura e estática	Otimização por Enxame de Partículas	1	127
El Aissaoui et al. (2018)	Felder-Silverman	Implícita, baseada em dados e dinâmica	Lógica Fuzzy	1	126
Fasihuddin, Skinner e Athauda (2017)	Felder-Silverman	Implícita, baseada em literatura e estática	-	1	83
Fatahi, Shabanali-Fami e Moradi (2017)	MBTI	Implícita, baseada em dados e estática	Mineração de Padrão Sequencial Generalizada	1	250

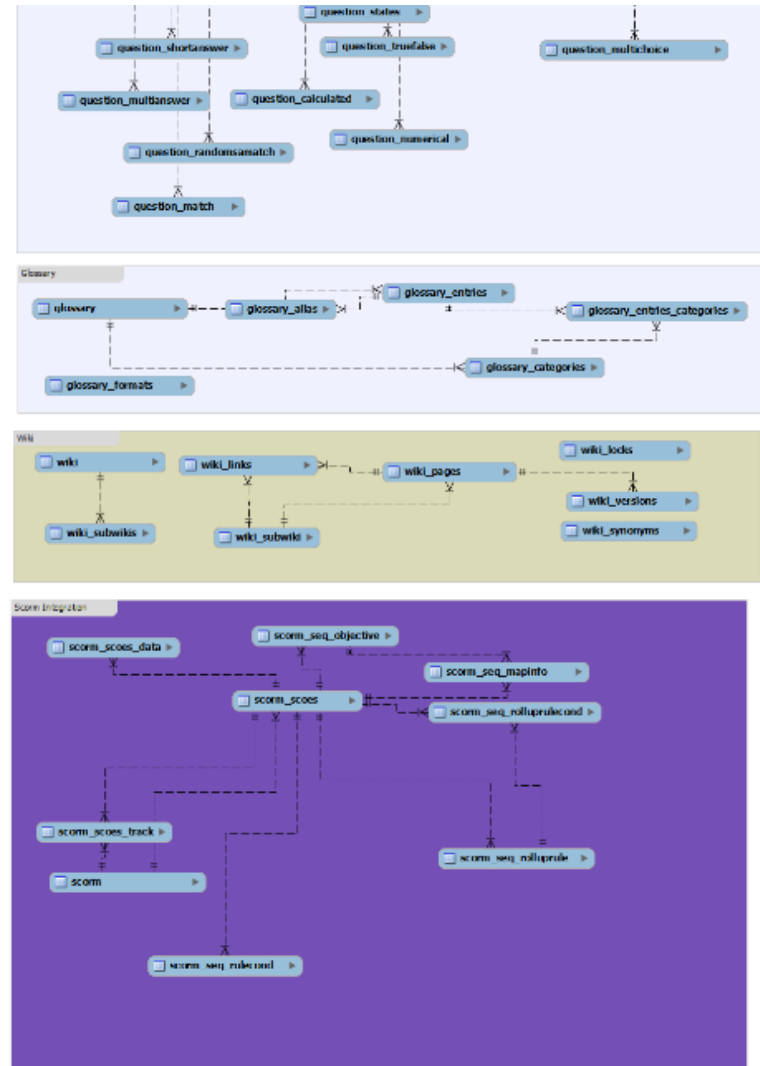
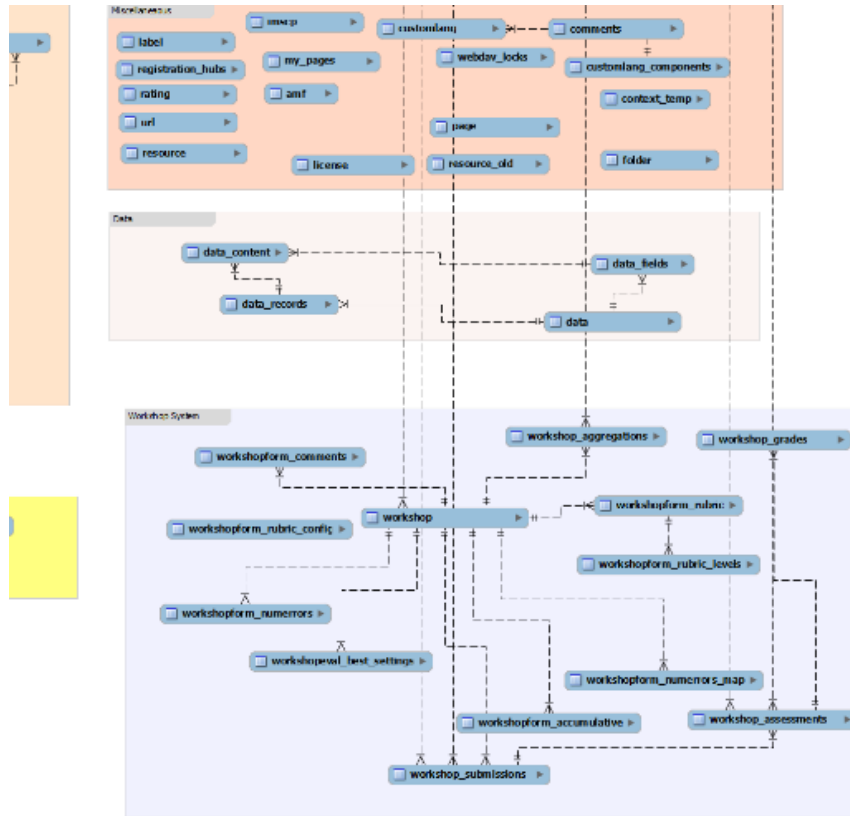
Ferreira et al. (2017)	Felder-Silverman	Implícita, baseada em literatura e estática	vários	1	180
Karagiannis e Satratzemi (2018)	Felder-Silverman	Implícita, baseada em dados e dinâmica	vários	1	139
Li e Abdul Rahman (2018)	Felder-Silverman	Implícita, baseada em literatura e dinâmica	Árvore Bayesiana	1	46
Lugo et al. (2015)	Felder-Silverman	Implícita, baseada em literatura e estática	Redes Bayesianas	1	134
Rajper, Saikh, Saikh e Mallah (2016)	Kolb	Implícita, baseada em dados e estática	Redes Bayesianas	vários	863
Sahid, Nugroho e Santosa (2017)	VARK	Implícita, mista e dinâmica	Modelo Oculto de Markov	1	60
Sena, Vivas, Assis e Pitangui (2016)	Felder-Silverman	Implícita, baseada em dados e estática	Modelo Oculto de Markov	-	-
Sheeba e Krishnan (2018)	Felder-Silverman	Implícita, baseada em literatura e estática	Árvore de Decisão	vários	300
Sweta e Lal (2015)	Felder-Silverman	Implícita, baseada em literatura e estática	Redes Bayesianas	1	40
Sweta e Lal (2016)	Felder-Silverman	Implícita, baseada em literatura e estática	Lógica Fuzzy	1	60
Sweta e Lal (2017)	Felder-Silverman	Implícita, baseada em	Lógica Fuzzy	1	120

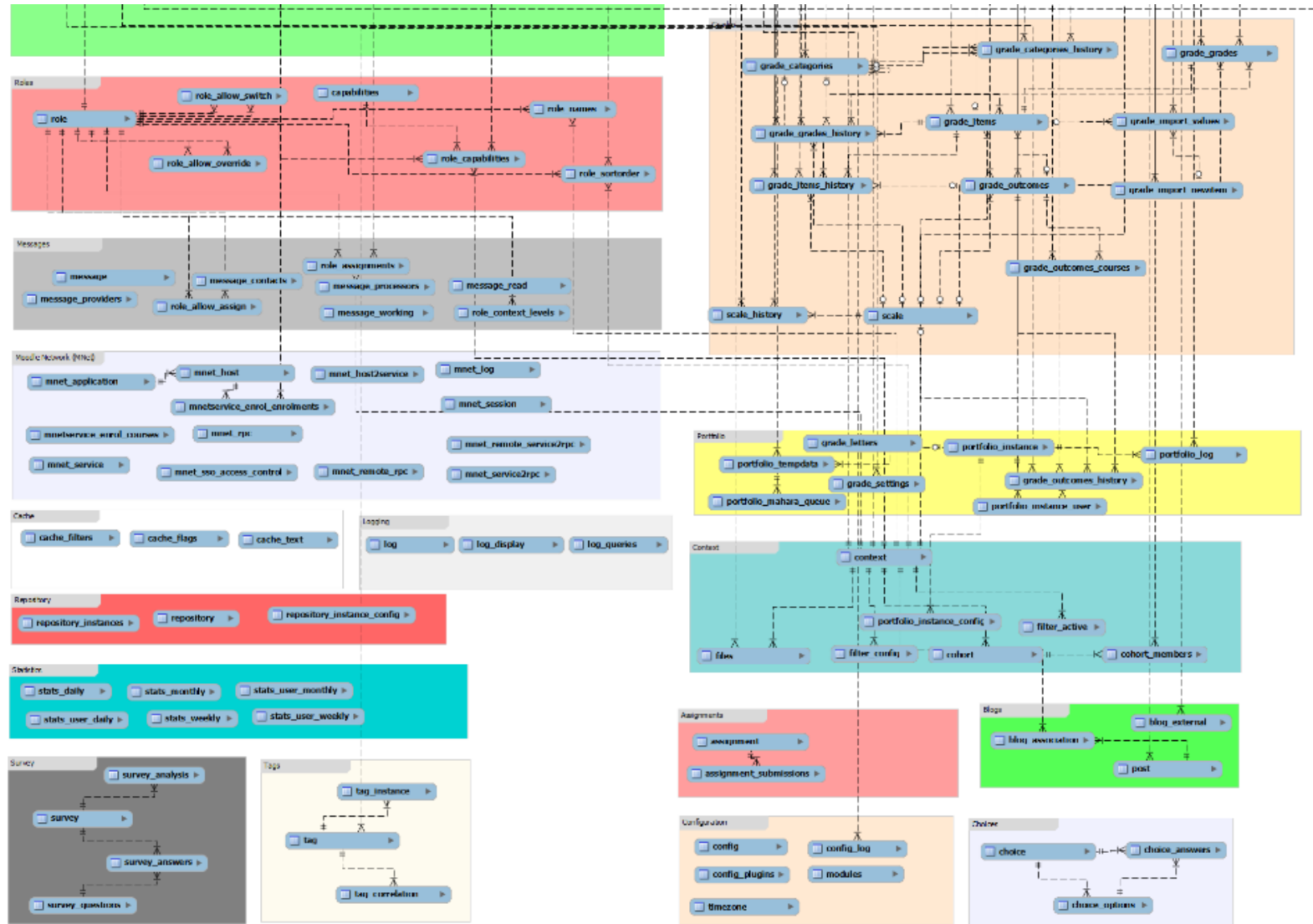
		literatura e estática			
Xiao e Rahman (2017)	Felder- Silverman	Implícita, baseada em dados e estática	-	1	33

APÊNDICE C – Estrutura relacional do banco de dados









APÊNDICE D – Exemplo de relatório apresentado aos respondentes

AGRADECEMOS SUA PARTICIPAÇÃO

Abaixo você pode verificar o resultado do seu teste de Estilo de Aprendizagem. Essas informações foram geradas a partir das suas respostas, com isso estatisticamente estes foram identificados como seus estilos de aprendizagem dominantes.

Processamento cognitivo

Pragmático: foca no que é concreto, onde não há dúvida. O mais importante para os pragmáticos é o quanto a informação é confiável, por isso, tendem a considerar apenas fatos e dados ao aprenderem. Sua aprendizagem é pautada em analisar e relacionar de forma clara e objetiva as informações que lhe são passadas.



Interação

Interage: aprende tanto com o material do curso quanto interagindo com colegas e tutores. Tirar suas dúvidas com outras pessoas acelera sua aprendizagem.

Interage mais ativamente do que passivamente: faz suas próprias perguntas, participando publicamente das discussões do curso.



Regulação

Autorregulado: controla sua própria aprendizagem tanto quanto ao ritmo quanto em relação à ordem do que aprende. Prefere cursos que lhe dão a autonomia para aprender da forma que considera melhor. Em geral, são organizados e metódicos.



APÊNDICE E – Análises descritivas da análise fatorial exploratória

Estatísticas Descritivas

	N	Mínimo	Máximo	Média	Desvio Padrão	Assimetria	Curtose
Logo antes das avaliações do curso, reviso os materiais de estudo.	303	1	5	4,33	,804	-1,246	1,677
Realizo todas as atividades do curso, mesmo as que não são pontuadas.	303	2	5	4,58	,586	-1,162	,877
Realizo um curso de cada vez.	303	2	5	4,47	,680	-1,033	,359
Acesso os materiais adicionais do curso.	303	1	5	3,55	,978	-,231	-,337
Defino minhas próprias metas para estudar os materiais do curso.	303	2	5	4,12	,764	-,345	-,817
Entrego com antecedência todas as tarefas do curso que têm prazo de entrega.	303	1	5	4,07	,765	-,470	,030
Leio todos os feedbacks que o tutor passa sobre as tarefas que realizei no curso.	303	1	5	3,12	,895	-,012	,190
Verifico minhas notas do curso assim que elas são lançadas.	303	1	5	3,79	1,020	-,404	-,797
Acesso os materiais do curso na ordem que são apresentados.	303	1	5	3,82	1,177	-,704	-,485
Acesso os materiais do curso várias vezes.	303	1	5	2,67	1,117	,413	-,411
Busco materiais além dos que são passados no curso.	303	1	5	2,80	1,004	,307	-,173
Divido os materiais de curso de forma a criar um ritmo de estudo que sigo até o final do curso.	303	1	5	2,57	1,077	,416	-,357
Realizo todos os exercícios do curso para praticar o que aprendi.	303	1	5	2,83	1,119	,208	-,635
Realizo apenas as atividades do curso que são pontuadas.	303	1	5	1,96	1,022	,979	,469

Aprendo os conteúdos do curso a partir de insights (ideias que surgem repentinamente).	303	1	5	3,17	1,157	-,049	-,726
Aprendo melhor sozinho os conteúdos do curso.	303	1	5	2,96	1,149	,058	-,743
Leio todas as perguntas, respostas e comentários que meus colegas e tutor postam no curso.	303	1	5	3,86	,816	-,544	,544
Leio todos os tópicos que são postados nos fóruns do curso.	303	1	5	3,23	1,187	-,425	-,663
Interajo com meus colegas de curso e com o tutor sobre assuntos que não estão diretamente ligados ao conteúdo.	303	1	5	4,23	,879	-1,021	,598
Posto minha dúvida quando tenho uma relacionada ao conteúdo do curso.	303	1	5	3,62	1,044	-,297	-,747
Quando tenho uma dúvida, pergunto a meus colegas de curso.	303	1	5	3,46	1,028	-,063	-,640
Interajo com o tutor durante o curso.	303	1	5	3,89	,896	-,474	-,247
Interajo com meus colegas de curso enquanto o realizo.	303	1	5	4,16	,763	-,765	,729
No curso, leio todas as respostas a postagens que faço.	303	2	5	4,46	,731	-1,276	1,159
Primeiro compreendo a estrutura geral do conteúdo do curso para depois estudá-lo em detalhes.	303	1	5	4,29	1,007	-1,449	1,481
Ao estudar os conteúdos do curso, busco conhecimentos fundamentados em fatos.	303	1	5	3,74	,922	-,420	-,264
Ao estudar os materiais do curso, foco em fatos e dados.	303	1	5	3,98	,922	-,592	-,381

Uso minha intuição (entender as coisas no momento, sem necessidade de realizar raciocínios complexos) para aprender o conteúdo do curso.	303	1	5	3,83	,929	-,308	-,710
Refliro sobre o conteúdo do curso para melhor compreendê-lo.	303	1	5	4,02	,900	-,593	-,233
Refliro sobre as possibilidades de aplicação no meu contexto profissional do que aprendi no curso.	303	1	5	3,76	1,001	-,445	-,434
Ao estudar o material do curso, foco em ideias para aprender.	303	1	5	3,76	,993	-,433	-,573
N válido (listwise)	303						

APÊNDICE F – Lista de informações retiradas do banco de dados

Informações do Banco de Dados		
<i>Informação</i>	<i>Categoria</i>	<i>Tipo</i>
Ordem em que acessa Recursos	Regulação	ordem
Ordem em que acessa Tarefas	Regulação	ordem
Recursos acessados antes de realizar a avaliação	Regulação	quantidade
Diferença de tempo entre primeira visualização de recurso e realização da prova	Regulação	tempo
Tempo Recursos durante a sessão	Regulação	tempo
Quantidade de recursos acessados em cada sessão	Regulação	quantidade
Quantidade de seções acessadas em cada sessão	Regulação	quantidade
Duração de cada sessão	Regulação	tempo
Diferença de tempo entre a realização das avaliações e o deadline das avaliações	Regulação	tempo
Diferença de tempo entre acesso a recursos ligados à avaliação e deadline da avaliação	Regulação	tempo
Duração de cada avaliação	Regulação	tempo
Quantidade de cursos anteriores finalizados	Regulação	quantidade
Quantidade de cursos anteriores evadidos	Regulação	quantidade
Relação entre cursos anteriores finalizados e evadidos	Regulação	proporção
Quantidade de cursos que esteve participando enquanto realizava o curso	Regulação	quantidade
Quantidade de sessões para finalizar o curso	Regulação	quantidade
Quantidade de sessões para finalizar os módulos	Regulação	quantidade
Tempo somado no curso	Regulação	tempo
Diferença de tempo entre a liberação das atividades e o acesso a elas	Regulação	tempo
Quantidade de acessos a recursos	Regulação	quantidade
Quantidade de postagens	Interação	quantidade
Relação entre a quantidade de postagens em fóruns e em discussões	Interação	razão
Postagens por categoria de Interação	Interação	quantidade
Quantidade de postagens que foram respondidas	Interação	quantidade
Relação entre postagens feitas e respondidas	Interação	razão
Relação entre postagens feitas e respondidas por outros estudantes	Interação	razão
Relação entre postagens feitas e respondidas pelo tutor	Interação	razão
Quantidade de postagens lidas	Interação	quantidade
Relação entre quantidade de postagens lidas e a serem lidas	Interação	razão
Relação entre postagens feitas e postagens lidas	Interação	razão
Quantidade de acessos a postagens que lê	Interação	quantidade
Diferença de tempo entre postagem e leitura	Interação	tempo
Diferença de tempo entre postagem e resposta	Interação	tempo
Diferença de tempo entre postagem do resultado da avaliação e acesso a ela	Regulação	tempo

Diferença de tempo entre postagem do feedback da avaliação e acesso a ele	Regulação	tempo
Duração de acesso a resultados de avaliações	Regulação	tempo
Duração de acesso a feedbacks de avaliações	Regulação	tempo
Quantidade de resultados de avaliações	Regulação	quantidade
Quantidade de acessos a feedbacks de avaliações	Regulação	quantidade

APÊNDICE G – Tipos de recursos presentes no banco de dados

Tipos de Recursos	
<i>Tipo</i>	<i>Recursos</i>
Avaliação	Tarefas
	Quiz
	Prova
Recurso	Módulos
	Seções
	Pastas
	Materiais
Interação	Fórum
	Discussão
	Chat

APÊNDICE H – Informações extraídas do banco de dados

Informações extraídas do banco de dados			
<i>Código</i>	<i>Descrição</i>	<i>Tipo de informação</i>	<i>Categoria</i>
UserXcourse	Entrada do estudante no curso	Básica	-
Course	Curso	Básica	-
User	Estudante	Básica	-
Discussoes_course	Quantidade de discussões no curso	Curso	Interação
Foruns_course	Quantidade de fóruns no curso	Curso	Interação
Posts_course	Quantidade de postagens feitas no curso	Curso	Interação
MedPostUserCourse	Média de postagens feitas por usuários no curso	Curso	Interação
MedDiscussUserPosta	Média de discussões em que os usuários fazem postagens no curso	Curso	Interação
MedForumUserCourse	Média de fóruns em que os usuários fazem postagens	Curso	Interação
MedLeituraForumCourse	Média de leituras de fóruns realizadas por usuários no curso	Curso	Interação
MedLeituraDiscussCourse	Média de leituras de discussões realizadas por usuários no curso	Curso	Interação
TarefasCurso	Quantidade de tarefas no curso	Curso	Avaliação
QuizCurso	Quantidade de quizzes no curso	Curso	Avaliação
ItensAvaliadosCurso	Quantidade de itens avaliados no curso	Curso	Avaliação
PostagensAluno	Quantidade de postagens do usuário	Estudante	Interação
PostsIniciado	Quantidade de postagens iniciadas pelo usuário	Estudante	Interação
PostsRespondendo	Quantidade de postagens feitas pelo usuário em resposta a outras postagens	Estudante	Interação
DiscussoesPostou	Quantidade de discussões onde o usuário postou	Estudante	Interação
ForunsPostou	Quantidade de fóruns onde o usuário postou	Estudante	Interação
PercentDiscussoesPostou	Percentual de discussões do curso onde o usuário postou	Estudante	Interação
PercentForumPostou	Percentual de fóruns do curso onde o usuário postou	Estudante	Interação

PercentTotalPosts	Percentual de posts do curso onde o usuário postou	Estudante	Interação
LeituraForum	Quantidade de leituras a fóruns realizadas pelo usuário	Estudante	Interação
LeituraDiscuss	Quantidade de leituras a discussões realizadas pelo usuário	Estudante	Interação
ForunsLidos	Quantidade de fóruns lidos pelo usuário	Estudante	Interação
DiscussoesLidas	Quantidade de discussões lidas pelo usuário	Estudante	Interação
CriacaoDiscuss	Quantidade de discussões criadas pelo usuário	Estudante	Interação
CriacaoPostsLog	Quantidade de posts criados (informação obtida a partir de tabela de logs)	Estudante	Interação
UpdatePostLog	Quantidade de updates feitos pelo usuário em seus próprios posts (informação obtida a partir de tabela de logs)	Estudante	Interação
TarefasRespondidas	Quantidade de tarefas respondidas pelo usuário	Estudante	Avaliação
PercentTarefasRespondidas	Percentual de tarefas respondidas pelo usuário	Estudante	Avaliação
MedSubmissoesTarefa	Média de submissões realizadas por tarefa	Estudante	Avaliação
NotaMedTarefas	Nota média obtida pelo usuário nas tarefas	Estudante	Avaliação
TarefasAvaliadas	Quantidade de tarefas avaliadas do usuário	Estudante	Avaliação
PercentTarefasAvaliadas	Percentual de tarefas avaliadas do usuário	Estudante	Avaliação
MedIntervaloDueDateTarefa	Diferença média de tempo entre o deadline da tarefa e a entrega dela pelo usuário	Estudante	Avaliação
MedIntervaloFromDateTarefa	Diferença média de tempo entre a postagem da tarefa e a entrega dela pelo usuário	Estudante	Avaliação
QuizAluno	Quantidade de quizzes respondidos pelo usuário	Estudante	Avaliação
MedTentativaQuiz	Média de tentativas realizadas por quiz pelo usuário	Estudante	Avaliação
NotaMedQuiz	Nota média obtida pelo usuário nos quizzes	Estudante	Avaliação

ItensAvalaidosRealizados	Quantidade de itens avaliados realizados pelo usuário	Estudante	Avaliação
PercentItensAvaliados	Percentual de itens avaliados realizados pelo usuário	Estudante	Avaliação
NotaMed	Nota média obtida pelo usuário nos itens avaliados	Estudante	Avaliação

APÊNDICE I – Indicadores criados a partir das informações extraídas do banco de dados

Indicadores Criados		
<i>Nome</i>	<i>Descrição</i>	<i>Categoria</i>
Tarefas Respondidas	Percentual de tarefas respondidas em relação ao total de tarefas do curso	Avaliação
Submissões por Tarefa	Média de submissões por tarefa	Avaliação
Nota Tarefas	Nota média nas tarefas	Avaliação
Tempo restante para entrega	Média da diferença de tempo entre o deadline da tarefa e a entrega dela pelo usuário	Avaliação
Tempo após divulgação	Média da diferença de tempo entre a postagem da tarefa e a entrega dela pelo usuário	Avaliação
Participação em Quiz	Quantidade de quizzes dos quais o estudante participou	Avaliação
Tentativas por Quiz	Quantidade média de tentativas por quiz	Avaliação
Nota Quiz	Nota média nos quizzes	Avaliação
Itens Avaliados Enviados	Percentual de itens avaliados enviados pelo estudante em relação ao total de itens avaliados no curso	Avaliação
Nota Final	Nota média nas atividades avaliadas	Avaliação
Posts	Quantidade de postagens do usuário	Interação
Posts Iniciados	Quantidade de postagens iniciadas pelo usuário	Interação
Posts Respondendo	Quantidade de postagens feitas pelo usuário em resposta a outras postagens	Interação
Posts Iniciados X Respondendo	Percentual de postagens iniciadas em relação ao total de postagens feitas	Interação
Discussões	Quantidade de discussões onde o usuário postou	Interação
Fóruns	Quantidade de fóruns onde o usuário postou	Interação
Discussões X Fóruns	Percentual de discussões onde o usuário postou em relação à soma das discussões e fóruns onde postou	Interação
Leitura de Fóruns	Quantidade de leituras de fóruns	Interação
Leitura de Discussões	Quantidade de leituras de discussões	Interação
Leitura de Discussões X Fóruns	Percentual de leituras a discussões em relação à soma das leituras a discussões e fóruns	Interação
Discussões X Leitura de Discussões	Razão entre discussões onde postou e quantidade de leituras de discussões	Interação
Fóruns X Leitura de Fóruns	Razão entre fóruns onde postou e quantidade de leituras de fóruns	Interação
Fóruns Lidos	Quantidade de fóruns lidos	Interação
Discussões Lidas	Quantidade de discussões lidas	Interação
Discussões Criadas	Quantidade de discussões criadas	Interação

Discussões Criadas X Postadas	Razão entre discussões criadas e discussões onde o usuário postou	Interação
Discussões Criadas X Lidas	Razão entre discussões criadas e leituras a discussões	Interação
Posts Atualizados	Quantidade de updates a posts realizados	Interação
Posts Atualizados X Posts	Razão entre quantidade de updates a posts realizados e posts realizados	Interação

**APÊNDICE J – Comunalidades da análise fatorial exploratória dos indicadores de
interação**

	Inicial	Extração
Posts	0,959	0,94
Posts Iniciados	0,933	0,862
Posts Respondendo	0,914	0,854
Posts Iniciados X Respondendo	0,888	0,812
Discussões	0,945	0,944
Fóruns	0,914	0,894
Leitura de Fóruns	0,885	0,803
Leitura de Discussões	0,889	0,813
Discussões X Leitura de Discussões	0,797	0,684
Fóruns X Leitura de Fóruns	0,811	0,567
Fóruns Lidos	0,651	0,539
Discussões Lidas	0,862	0,812
Discussões Criadas	0,936	0,918
Discussões Criadas X Postadas	0,974	0,909
Discussões Criadas X Lidas	0,978	0,929
Posts Atualizados	0,861	0,998
Posts Atualizados X Posts	0,851	0,845

APÊNDICE K – Scree plot da análise fatorial exploratória dos indicadores de interação