

# O Efeito das Características dos Alunos em Modelos de Dinâmica de Afetos e Detecção Livre de Sensores

Felipe de Moraes<sup>1</sup>, Patricia A. Jaques<sup>2</sup>

<sup>1</sup>PPGCA - Universidade do Vale do Rio dos Sinos (UNISINOS)

<sup>2</sup>PPGInf - Universidade Federal do Paraná (UFPR)  
PPGC - Universidade Federal de Pelotas (UFPEL)

felipedemoraism@hotmail.com, patricia@inf.ufpr.br

**Abstract.** *This study focuses on investigating affect dynamics among Brazilian students while developing sensor-free emotion detection models that consider additional student characteristics. The research addresses the problem of understanding how different factors related to personality, motivation, gender, behavior, and emotion duration influence affect dynamics and emotion detection for engagement, confusion, boredom, and frustration in the context of Intelligent Tutoring Systems (ITS). The research method involves data collection using the ITS PAT2Math and the EmAP-ML protocol for emotion annotation. The main results indicate that the affect dynamics model in the Brazilian context is similar to more recent models found in the literature. Additionally, it was possible to identify that student characteristics also influence transitions between student emotions. Furthermore, it was found that using student characteristics as additional data sources improves the performance of sensor-free emotion detectors.*

**Resumo.** *Este estudo foca-se nas emoções envolvidas na aprendizagem, com o objetivo de investigar a dinâmica afetiva entre estudantes brasileiros e desenvolver modelos de detecção de emoções livres de sensores, considerando características adicionais dos alunos. A pesquisa busca entender como diversos fatores, tais como personalidade, motivação, sexo, comportamentos e a duração das emoções, influenciam na dinâmica afetiva e na detecção das emoções engajamento, confusão, tédio e frustração no âmbito dos Sistemas de Tutoria Inteligentes (STI). A metodologia empregada envolve a coleta de dados por meio do STI PAT2Math, e a utilização do protocolo EmAP-ML para a anotação das emoções. Os resultados principais sugerem que o modelo de dinâmica afetiva aplicado ao contexto brasileiro é análogo aos modelos mais recentes disponíveis na literatura. Ademais, foi possível constatar que as características individuais dos estudantes influenciam nas transições entre suas emoções. Observou-se ainda que a inclusão dessas características como fontes de dados adicionais otimiza o desempenho dos detectores de emoção sem sensores.*

## 1. Introdução

Emoções podem impactar tanto positiva [Fredrickson 1998] quanto negativamente [Graesser and D’Mello 2011] o aprendizado, influenciando atenção, motivação, cognição e autorregulação dos alunos [Pekrun 2014]. Em Ambientes Computacionais de Aprendizagem (ACA), emoções como confusão, frustração, tédio e engajamento são comuns [Calvo and D’Mello 2010], conhecidas como emoções acadêmicas [Pekrun 2016].

Diversos fatores influenciam emoções durante o aprendizado. A **personalidade** pode determinar que emoções são vivenciadas [Tong 2010] e sua duração [Goldoni et al. 2022]. A **motivação** direciona atenção, afetando autoavaliações e cognições [Linnenbrink-Garcia and Barger 2014]. O **sexo**<sup>1</sup> pode influenciar a quantidade de emoções vivenciadas [Frenzel et al. 2007]. **Comportamentos** redirecionam emoções no ACA [Sabourin et al. 2011], influenciando a transição de emoções negativas para positivas [Baker et al. 2011a]. Emoções podem afetar emoções subsequentes [Bosch and D’Mello 2017, Sinclair et al. 2018, Andres et al. 2019].

O modelo proposto por [D’Mello and Graesser 2012] afirma que emoções não ocorrem aleatoriamente e que estados emocionais negativos requerem intervenções pedagógicas para regular. Contudo, esse modelo pode não ser generalizável para diferentes culturas [Karumbaiah et al. 2021]. Não foram identificadas pesquisas que investigaram a influência de características dos estudantes relacionadas à personalidade, motivação, sexo e comportamentos de aprendizagem nos modelos de dinâmica afetiva.

Considerando o impacto das emoções na aprendizagem, ACAs devem detectar emoções dos alunos [D’Mello and Graesser 2012, Arroyo et al. 2014] e adaptar suas estratégias pedagógicas [Woolf et al. 2009, Conati 2011, Arroyo et al. 2014]. Muitos métodos de detecção de emoções se baseiam em dados de sensores, o que aumenta custos e intrusões [Calvo and D’Mello 2010], tornando inviável a aplicação em escolas. Um método mais viável seria utilizar a interação dos estudantes com o sistema [Harley 2016], conhecido como detecção de afeto livre de sensores [Baker et al. 2012].

Ao analisar o estado da arte em detecção de emoções sem sensores, foi possível identificar um conjunto de trabalhos apresentando resultados muito significativos, mas que ainda não atingiram uma precisão ideal quando comparados ao desempenho humano [Baker et al. 2010a, D’mello et al. 2008, Paquette et al. 2014, Jiang et al. 2018, Paquette et al. 2016, Henderson et al. 2020, Hutt et al. 2019]. Além disso, nenhuma pesquisa estudou a influência das características dos alunos, como personalidade, motivação, sexo, comportamentos e dinâmica afetiva nos detectores de emoção sem sensor.

Portanto, este trabalho visa preencher essas lacunas através de dois estudos em emoções em ACAs. O Estudo 1 desenvolve um modelo de dinâmica afetiva para estudantes brasileiros, considerando suas características pessoais. Ao analisar diferenças nos modelos de dinâmica afetiva baseados nesses dados, é possível entender melhor o impacto das características dos alunos em suas emoções. Já o Estudo 2 objetiva desenvolver um modelo de detecção afetiva sem sensores, considerando as características pessoais dos alunos. Foram utilizados dados de interação de estudantes do sétimo ano usando o PAT2Math [Jaques et al. 2013], um Sistema Tutor Inteligente que fornece orientação durante a resolução de equações de primeiro grau.

## 2. Estado da Arte

Como o objetivo deste trabalho é duplo, estudar a influência das características dos estudantes em modelos de dinâmica de afeto e em detectores livre de sensores, foi realizada uma revisão do estado da arte nesses dois temas. Inicialmente, realizou-se uma análise

---

<sup>1</sup>Utiliza-se o termo sexo para remeter as categorias inatas da biologia humana, feminino e masculino, diferindo do gênero que remete aos papéis sociais da mulher e do homem na sociedade [Oakley 2016].

da literatura sobre a dinâmica de afetos dos alunos com base em uma busca nas seguintes bases de dados: nacional (base de dados SBC) e internacional (ACM, IEEE, Web of Science, Springer, Scopus, Science Direct e Society of Educational Data Mining). O modelo teórico proposto por [D’Mello and Graesser 2012] sugere que os alunos engajados enfrentam desafios durante o aprendizado, experimentando confusão devido a um desequilíbrio cognitivo. Se essa confusão for adequadamente regulada, os alunos se engajam novamente e aprendem com essa experiência. Caso contrário, a confusão se transforma em frustração e tédio, deixando o aluno bloqueado e sem saber como proceder. Esse modelo teórico tem servido de base para diversas pesquisas. No entanto, [Karumbaiah et al. 2021] identificou que os modelos de dinâmica de afetos são dependentes do contexto dos alunos, especialmente em relação ao país de origem.

Posteriormente, foram buscados trabalhos que consideraram as características individuais dos alunos no desenvolvimento dos modelos. No estudo de [Baker et al. 2010b], os autores investigaram a influência das emoções e comportamentos dos alunos por meio da métrica L. Os resultados sugerem que é necessário levar em consideração o tédio persistente para evitar que os alunos entrem em um “ciclo vicioso”. De forma semelhante, [Baker et al. 2011b] analisaram as emoções que precedem, acompanham e sucedem os comportamentos dos alunos usando um ACA de química. Os autores destacam que o comportamento “gaming the system” alivia a confusão, mas prediz o tédio posteriormente. Não foi identificado nenhum outro estudo que considere as características individuais dos alunos no desenvolvimento de modelos de dinâmica de afetos.

No tema de detectores de emoção sem o uso de sensores, realizou-se uma busca por trabalhos nas mesmas bases de dados mencionadas anteriormente, utilizando os termos “emoção”, “detecção” e “ambientes de aprendizagem”. A partir de um conjunto inicial de 369 artigos, apenas artigos completos com foco na detecção automática de emoções por meio de dados de logs foram selecionados [Pardos et al. 2014, Ocumpaugh et al. 2014, Paquette et al. 2014, Paquette et al. 2016, Botelho et al. 2017, Jiang et al. 2018, Hutt et al. 2019, Henderson et al. 2020]. Os estudos relacionados priorizaram principalmente a detecção de quatro emoções: confusão, tédio, engajamento e frustração, embora algumas exceções tenham considerado emoções adicionais. Esses estudos compartilham uma série de métodos comuns. Primeiramente, a coleta de rótulos de emoção foi realizada utilizando o protocolo de anotação de emoções BROMP [Ocumpaugh 2015], com exceção do estudo de [Hutt et al. 2019], que optou pelo método de autoavaliação. O protocolo BROMP permite a coleta de emoções *online*, em sala de aula, enquanto os alunos utilizam o ambiente educacional. Em seguida, um conjunto de algoritmos foi empregado<sup>2</sup>, sendo cada um testado para cada emoção considerada. O algoritmo com melhor desempenho foi selecionado para cada emoção. Esses estudos também aplicaram a técnica de validação cruzada, com variação entre 5 ou 10 dobras, no nível do estudante. Além disso, foi realizada uma fase de engenharia de características ou aplicada uma abordagem de aprendizado de máquina para calcular as características. Para cada emoção considerada, foi desenvolvido um modelo de detecção, transformando os rótulos de emoção em classes binárias e utilizando a abordagem um-contratodos para treinar os modelos, onde cada modelo é treinado para identificar a presença ou ausência de

---

<sup>2</sup>Lista de algoritmos empregados: JRip, Logistic Regression, J48, Naïve Bayes, REP-Trees, KStar, Step Regression, Multi-Layer Perceptron, Support Vector Machine, and Long-Short Term Memory network.

cada emoção. Por fim, as métricas de avaliação Kappa de Cohen e AUC foram aplicadas, com exceção dos estudos de [Henderson et al. 2020] e [Hutt et al. 2019], que utilizaram apenas a correlação de Spearman rho e F1-score, respectivamente.

A revisão da literatura sobre modelos de dinâmica de afetos e detecção de emoções sem o uso de sensores em ambientes de aprendizagem revela algumas lacunas. Em relação aos modelos de dinâmica de afetos, percebe-se uma carência de modelos desenvolvidos especificamente para o contexto brasileiro. Além disso, nenhum dos modelos existentes considera as características individuais dos alunos, como personalidade, sexo, motivação e duração das emoções. No que diz respeito aos estudos de detecção de emoções sem o uso de sensores, esses também não levam em conta as características individuais dos alunos para aprimorar o desempenho do modelo. Apenas dois estudos consideraram o comportamento do aluno, porém não analisaram o impacto dessa informação na detecção das emoções acadêmicas. Essas lacunas serviram de motivação para o desenvolvimento dos dois estudos que compõem esta pesquisa.

### 3. Método

Realizou-se uma coleta de dados que serviu como base para a elaboração de ambos os estudos. Os dados foram obtidos com alunos do sétimo ano de uma escola particular do RS, utilizando um Sistema de Tutoria Inteligente (STI) orientado a passos, o PAT2Math, que tem contribuído para o aprendizado de álgebra para centenas de alunos no Brasil [Jaques et al. 2013]. O STI foi empregado por 55 alunos, distribuídos em 10 sessões de aproximadamente 40 minutos cada. Esse processo resultou na produção e no armazenamento de 386.015 registros de interação, denominados logs. Cada log incorporou um total de 329 características, incluindo dados sobre cliques, teclas pressionadas, dicas, passos, operações, entre outros. Adicionalmente, ao longo dessa coleta de dados, os alunos responderam a questionários de personalidade [Barbosa 2009] e motivação [Gouveia et al. 2008]. Assim, foram recolhidos os dados referentes a traços de personalidade e metas de realização dos 55 estudantes que concordaram em participar deste estudo e assinaram o Termo de Consentimento Livre e Esclarecido (TCLE)<sup>3</sup>.

As emoções e comportamentos dos alunos foram obtidas através do protocolo EmAP-ML. O protocolo EmAP-ML permite a coleta de emoções e comportamentos dos alunos através da análise de vídeos dos alunos utilizando o STI. Neste protocolo, três anotadores humanos treinados anotaram as emoções de confusão, frustração, tédio, engajamento e “outro” (rótulo para todas as emoções identificadas que não as citadas) ao analisar as gravações dos rostos dos alunos com áudio e seus vídeos de tela do computador [Morais et al. 2019]. Um total de 30 vídeos, provenientes das dez sessões de coleta de dados, foi analisado. Para uma seleção equilibrada de vídeos entre as sessões, três vídeos de cada sessão foram selecionados de forma aleatória entre os alunos, mas o mesmo aluno não foi selecionado mais de uma vez. De acordo com o estabelecido na fase de treinamento dos codificadores [Morais et al. 2019], a duração da sessão de anotação é de cinco minutos, resultando em 150 minutos de vídeos anotados. Como o tamanho do clipe (cada trecho de vídeo anotado) é de cinco segundos, foram realizadas 1800 anotações. Contudo, uma vez que os codificadores podem anotar até duas emoções ou comportamentos por clipe, o total de rótulos de emoções chegou a 2099 e 2059 rótulos de comportamentos.

---

<sup>3</sup>TCLE sob o CAAE 30225220.0.1001.5344 na plataforma Brasil.

No primeiro estudo (E1), os rótulos de emoções dos alunos foram empregados para calcular a estatística L [D’Mello et al. 2007], com o objetivo de identificar a probabilidade de transição entre duas emoções, isto é, inferir se uma transição é significativamente mais provável do que o acaso. Nesta etapa, foi aplicado o procedimento estatístico sugerido por [Karumbaiah et al. 2021], considerando a métrica L ajustada. O modelo geral de dinâmica de afetos foi desenvolvido de acordo com as probabilidades de transição entre cada emoção. Posteriormente, os dados foram segmentados de acordo com as diferentes características dos alunos para verificar como estas influenciam individualmente nos modelos de dinâmica afetiva. Ao todo, foram construídos 17 modelos de dinâmica de afetos dos alunos. Seguindo a metodologia de [Karumbaiah et al. 2021], também foram realizados testes t para calcular a significância estatística das transições e a aplicação do procedimento de correção post-hoc de Benjamini-Hochberg (BH) para controlar resultados falsos-positivos. Assim, somente dados estatisticamente significativos foram considerados para a comparação dos resultados.

No segundo estudo (E2), foi desenvolvido um modelo referência de detecção de emoções sem sensores para cada emoção, no qual apenas os dados de log, ou seja, as características relacionadas ao comportamento observável dos alunos, foram utilizadas. Depois, modelos hipotéticos foram desenvolvidos utilizando os mesmos dados de logs com a adição de características dos traços de personalidade, orientação motivacional, sexo, comportamentos e informações sobre dinâmica afetiva. Considerando que as características podem interagir entre si, foi criado um modelo que incorpora as mesmas características do modelo de referência junto com todas as cinco características pessoais dos alunos. Ao todo, 28 modelos de detecção de emoções livres de sensores foram desenvolvidos.

O mesmo método foi seguido para criar os modelos de referência e os modelos hipotéticos de detecção das emoções. Construiu-se um *pipeline* de aprendizado de máquina que executa cada etapa sequencialmente e evita vazamento dos dados de treino nas fases de teste. As etapas implementadas no pipeline incluíram padronização de dados (*standard scalar*), seleção de características (*constant elimination, select 100 best F-values, 10 best forward selected*), seleção do modelo (*Logistic Regression, Decision Tree, Random Forest, Naive Bayes, Support Vector Machine, K-Nearest Neighbors, Stochastic Gradient Descent, Multi-Layer Perceptron, AdaBoost, e XGBoost*), reamostragem de dados (SMOTE e Tomek-links) e otimização de hiperparâmetros (Grid Search com validação cruzada). O conjunto de dados foi dividido em treinamento, validação e teste. O treinamento e a validação foram divididos automaticamente por meio de validação cruzada estratificada de 10 dobras a nível de log. Após o treinamento, os resultados de desempenho dos modelos foram comparados com o conjunto de teste de acordo com o Kappa de Cohen. Por fim, o teste estatístico de McNemar foi aplicado para verificar a diferença entre os modelos de referência e os modelos hipotéticos para cada emoção.

## **4. Resultados e Discussões**

Esta seção delinea os principais resultados de cada estudo, separados entre dinâmica de afetos (E1) e detecção de emoções sem sensores (E2).

### **4.1. Dinâmica de Afetos (E1)**

O E1 tinha como objetivo realizar uma análise da dinâmica de afetos no contexto brasileiro e verificar a similaridade desses resultados com a literatura. Posteriormente, esse

modelo geral, incluindo todos os alunos participantes na coleta de dados, foi comparado com modelos que consideravam as características dos alunos, como personalidade, motivação, sexo, comportamentos e duração das emoções.

No que se refere ao primeiro objetivo, identificou-se que existe um modelo de dinâmica de afetos com transições significativas entre as emoções de aprendizagem. Verificou-se a presença de dois ciclos mais prováveis. O primeiro ocorre entre os estados de engajamento e confusão, sugerindo que é comum que estudantes engajados se tornem confusos ( $L = 0,28$ ) e, ainda mais provavelmente, que estudantes confusos voltem a se engajar ( $L = 0,50$ ). Esse ciclo já foi relatado em diversos estudos, como [D’Mello and Graesser 2012] e [Karumbaiah et al. 2021], para alunos dos Estados Unidos. Embora não se tenha identificado transições mais propensas para sair da confusão, [D’Mello et al. 2014] descrevem que a confusão é benéfica para a aprendizagem quando resolvida, mas pode ser prejudicial caso contrário<sup>4</sup>.

Posteriormente, um modelo de dinâmica de afetos foi desenvolvido para cada sexo, separando os dados de meninos e meninas, a fim de comparar o modelo geral com os modelos que consideram o sexo e entre modelos de sexos diferentes. A comparação dos modelos baseados no sexo com o modelo geral demonstra que as transições diferem entre esses dois modelos. Ao calcular as probabilidades de transição considerando o sexo dos alunos, a principal descoberta está na transição significativa do tédio para o engajamento encontrada apenas em estudantes do sexo feminino ( $L = 0,49$ ). Mesmo identificando diferentes transições, as probabilidades de transição se mantêm na mesma direção (positiva ou negativa), independentemente do sexo do aluno. Assim, em vez de influenciar os modelos de dinâmica afetiva, acredita-se que o sexo dos alunos auxilie na identificação de transições significativas entre as emoções<sup>5</sup>.

A coleta de emoções dos estudantes foi realizada seguindo o protocolo EmAP-ML [Morais et al. 2019], o qual permitiu a captura do tempo que os alunos dedicavam a cada emoção antes de transitar para outra. Com base nesses dados, elaboraram-se dois modelos que levam em conta o tempo de duração das emoções: um considerando apenas as transições em que a duração da emoção foi inferior à mediana e outro considerando apenas as transições em que essa duração superou a mediana para cada emoção. Assim, foi viável comparar o modelo geral com os modelos de duração de emoção, bem como realizar comparações entre os próprios modelos.

Os resultados indicaram que as transições entre emoções dependem da duração. As transições do tédio para o engajamento ( $L = 0,57$ ), tédio para o estado “outro” ( $L = -0,18$ ), engajamento para o tédio ( $L = -0,02$ ) e confusão para o estado “outro” ( $L = -0,09$ ) foram significativas apenas quando os estudantes permaneceram um tempo inferior à mediana em cada uma das emoções de origem. Em contraste, a transição do engajamento para o estado “outro” ( $L = 0,24$ ) foi significativa apenas quando os alunos se mantiveram engajados por um tempo superior à mediana<sup>6</sup>.

No que tange à personalidade, desenvolveram-se modelos de dinâmica de afetos considerando os traços de personalidade dos alunos, em acordo com a teoria dos cinco

---

<sup>4</sup>Os resultados deste modelo foram descritos de forma mais detalhada em [Morais and Jaques 2021].

<sup>5</sup>Os resultados desta comparação foram descritos em [Morais and Jaques 2022].

<sup>6</sup>Esses dados são detalhados em [Morais and Jaques 2023].

grandes fatores [John and Srivastava 1999]. Os traços de personalidade foram obtidos através de um questionário validado para adolescentes brasileiros [Barbosa 2009]. Posteriormente, calculou-se o T-score de cada traço para cada aluno, a fim de identificar a intensidade (mais forte ou mais fraca) de cada traço por aluno [de Araújo et al. 2021]. Finalmente, foram construídos dois modelos por traço, levando em conta os dados dos alunos com presença mais forte e mais fraca de cada traço. Os resultados revelaram que cada modelo, considerando a presença mais forte e mais fraca de cada traço de personalidade, apresenta diferenças significativas nas transições entre emoções. Por exemplo, somente alunos com traço de amabilidade fraco apresentam a transição de engajamento para confusão como significativa. Somente alunos com traço de extroversão forte apresentaram a transição de tédio e confusão para engajamento como significativo. Alunos com o traço de conscienciosidade forte não tem nenhuma transição significativa saindo de engajamento. Somente alunos com o traço abertura forte apresentam a transição de tédio para engajamento como significativo. E, alunos com forte traço de neuroticismo apresentam poucas transições significativas entre emoções, indicando uma falta de sistematicidade.

Modelos de dinâmica de afetos foram desenvolvidos considerando a orientação motivacional dos alunos, conforme a teoria das metas de realização, que destaca duas razões para o engajamento dos estudantes em comportamentos de realização: as metas de domínio e de desempenho [Dweck and Leggett 1988]. Informações sobre a orientação motivacional dos alunos foram coletadas através de um questionário validado [Gouveia et al. 2008] e foi calculado um índice para identificar a orientação motivacional predominante em cada estudante. Contudo, devido à falta de dados, foram desenvolvidos apenas os modelos para as metas de busca do domínio e de evitação do domínio, não incluindo as metas de desempenho.

Os resultados mostraram que alunos com orientação de evitação do domínio possuem baixa probabilidade de transição de outros estados para o tédio ( $L = -0,01$ ). Por outro lado, alunos com orientação de busca do domínio têm baixa probabilidade de transição da frustração para o tédio ( $L = -0,08$ ), da confusão para o tédio ( $L = -0,02$ ), de outros estados para confusão ( $L = -0,10$ ) e alta probabilidade de transição da confusão para o engajamento ( $L = 0,51$ ) e do tédio para o engajamento ( $L = 0,42$ ). Ao comparar esses resultados com o modelo geral, observa-se que alunos com orientações de evitação e de busca do domínio apresentam transições semelhantes às encontradas no modelo geral.

Os comportamentos *on-task* (trabalhando na tarefa), *on-task conversation* (trabalhando na tarefa e conversando sobre ela), *on-task resource* (trabalhando na tarefa usando o caderno), *on-system* (somente mexendo no sistema) e *off-task* (sem trabalhar na tarefa) foram coletados juntamente com as emoções durante a fase de anotação pelos codificadores, seguindo o protocolo EmAP-ML [Morais et al. 2019]. Isso possibilitou que as emoções e comportamentos fossem sincronizados, permitindo calcular a relação entre eles usando a métrica  $L$ , conforme realizado em [Baker et al. 2011b]. Devido à simultaneidade das emoções e comportamentos, foi possível investigar se eles poderiam influenciar um ao outro, calculando a probabilidade de: uma emoção (anterior) prever o início ou continuação de um comportamento (seguinte), um comportamento anterior prever a emoção seguinte, e a co-ocorrência entre cada par de emoção e comportamento.

Os resultados demonstraram que todas as emoções preveem se os alunos se tornarão *on-task*, exceto a frustração, que é significativa apenas para alunos do sexo mas-

culino ( $L = 0,03$ ). De todas as emoções, apenas o engajamento prevê comportamentos diferentes de *on-task* (considerando apenas probabilidade positiva), ou seja, *off-task* ( $L = 0,12$ ) e *on-system* ( $L = 0,17$ ) (apenas para alunos do sexo masculino). Todos os comportamentos preveem o engajamento, exceto *off-task* ( $L = 0,53$ ) e *on-task conversation* ( $L = 0,43$ ), que são significativas apenas para alunas do sexo feminino. Apenas o comportamento *on-task* prevê emoções diferentes do engajamento (considerando apenas probabilidade positiva), como a confusão ( $L = 0,17$ ) e o estado “outro” ( $L = 0,20$ ). Além disso, ao considerar a co-ocorrência de emoções e comportamentos, *on-task* ocorre simultaneamente com a confusão ( $L = 0,24$ ) e com o engajamento ( $L = 0,64$ ). *On-system* ocorre com o engajamento para alunos do sexo masculino ( $L = 0,28$ ), com o tédio para alunas do sexo feminino ( $L = 0,22$ ) e com o estado “outro” para todos ( $L = 0,59$ ). *On-task resource* ocorre fortemente com o engajamento ( $L = 0,82$ ). *Off-task* ocorre fortemente com o estado “outro” ( $L = 0,76$ ). E *on-task conversation* ocorre com a confusão ( $L = 0,66$ ).

Foi observado que o modelo geral de dinâmica afetiva deste trabalho, em um contexto brasileiro, é semelhante ao modelo apresentado por [Karumbaiah et al. 2021], considerando alunos dos EUA, e difere de [D’Mello and Graesser 2012] na maioria das transições. A similaridade com [Karumbaiah et al. 2021] sugere a possibilidade de existência de um modelo geral de dinâmica afetiva. No entanto, evidências estatísticas encontradas indicam que características adicionais dos alunos, como personalidade, motivação, sexo, comportamentos e dados da duração das emoções, podem influenciar as emoções que os alunos experienciam durante a aprendizagem com um STI. Assim, os resultados sugerem que o modelo geral de dinâmica afetiva pode depender das informações pessoais dos alunos e da duração das emoções.

É relevante destacar a diferença entre ACAs. Ao comparar as emoções dos alunos em um ambiente baseado em jogos e um STI, [Rodrigo et al. 2008] sugerem que uma combinação de características de jogos e STIs seria mais eficaz. Neste trabalho, os alunos utilizaram o PAT2Math com uma interface gamificada. Essa estratégia foi desenvolvida anteriormente para evitar comportamentos dos alunos prejudiciais à aprendizagem, como burlar o sistema e recusa de ajuda [Azevedo et al. 2018]. Constatou-se que os estudantes utilizam o comportamento *on-system*, associado à interação com o sistema (não resolvendo tarefas) ou aos elementos de gamificação, para aliviar o tédio.

#### **4.2. Detecção de Emoções Livre de Sensores (E2)**

No experimento E2, buscou-se identificar se características dos estudantes, tais como personalidade, motivação, sexo, comportamentos e dinâmica emocional, poderiam ter um impacto positivo no desempenho dos modelos de detecção de emoções sem sensores na identificação de confusão, tédio, frustração e engajamento. Para a realização do E2, foram elaborados dois tipos de detectores de emoções para validar as hipóteses: modelos de referência e modelos hipotéticos. Os modelos de referência foram construídos considerando apenas os registros de interação (logs), enquanto os modelos hipotéticos utilizaram os logs e informações adicionais dos alunos conforme cada hipótese, seguindo a mesma metodologia de desenvolvimento<sup>7</sup>.

Ao examinar os modelos de detecção levando em consideração as características

---

<sup>7</sup>Os resultados dos modelos de referência e hipotéticos mais bem-sucedidos para cada emoção são apresentados em <https://bit.ly/3JHLARW>.



de personalidade, motivação e sexo dos alunos, observa-se que não houve diferença significativa entre os modelos hipotéticos e os modelos de referência para todas as emoções acadêmicas. Mesmo que tais características não tenham aprimorado a detecção de emoções, os detectores de confusão e tédio selecionaram características da personalidade entre as 10 mais relevantes, quando comparados com os respectivos modelos de referência. Notou-se também que o detector de engajamento listou o sexo do estudante entre as 10 características mais significativas.

Ao analisar os modelos de detecção tendo em vista os comportamentos dos alunos durante o aprendizado, observa-se que três (para engajamento, frustração e tédio) dos quatro detectores de emoções acadêmicas apresentaram melhora significativa entre os modelos hipotéticos e os de referência. O detector de confusão foi o único que não mostrou diferença significativa em relação ao modelo de referência correspondente. Entretanto, esse detector selecionou a característica dos comportamentos entre as 10 mais relevantes. Os detectores de engajamento, frustração e tédio, que consideraram os comportamentos de aprendizagem, melhoraram seus desempenhos de um Kappa de 0,75 para 0,80, 0,78 para 0,87, e 0,75 para 0,81, respectivamente, com significância estatística pelo teste de McNemar. De acordo com as características selecionadas no modelo hipotético, os comportamentos *on-task*, *on-task conversation*, e *on-task resource* foram selecionados entre os 10 mais importantes para engajamento. O comportamento *on-system* foi selecionado entre as 10 mais importantes para frustração. Os comportamentos *on-task* e *off-task* foram selecionados entre os 10 mais importantes para o tédio.

Ao examinar os modelos de detecção que consideram as características sobre a dinâmica de afetos dos alunos durante o aprendizado, observa-se que apenas o detector de frustração apresentou melhora significativa entre os modelos hipotéticos e os modelos de referência. Este detector de frustração, considerando a dinâmica de afeto, melhorou o desempenho dos detectores de um Kappa de 0,78 para 0,88, com significância estatística pelo teste de McNemar. De acordo com as características selecionadas no modelo hipotético, as transições confusão → frustração, tédio → frustração e outro → tédio foram selecionadas entre as 10 mais relevantes.

Por fim, modelos foram desenvolvidos para a detecção das emoções considerando todas as características dos alunos de uma só vez. Ao analisar os modelos de detecção que consideram características de todas as fontes de dados, observa-se que os quatro detectores de emoção apresentaram uma diferença significativa entre os modelos hipotéticos e os de referência. No entanto, ao avaliar os resultados, apesar de todos os detectores terem mostrado melhora significativa quando todas as fontes de dados foram consideradas, os detectores de engajamento, confusão e frustração selecionaram características relacionadas a apenas uma fonte de dados. Portanto, a combinação de fontes de dados não foi considerada como influenciadora positiva. Por outro lado, o detector de tédio, ao considerar as características de todas as fontes de dados, melhorou o desempenho do detector de um Kappa de 0,75 para 0,81, com significância estatística pelo teste de McNemar. De acordo com as características selecionadas no modelo hipotético, os comportamentos *on-task* e *off-task*, e a meta de busca do domínio (motivação), foram selecionados entre as 10 mais importantes. Este foi o único modelo, entre os quatro detectores considerando todas as fontes de dados, que considerou múltiplas fontes - comportamentos e motivação - e obteve uma diferença significativa em relação ao modelo de referência.

## 5. Conclusão

Este trabalho teve como objetivo investigar a influência de informações adicionais sobre personalidade, motivação, sexo e comportamentos de aprendizagem do aluno nos modelos de dinâmica de afetos (E1) e sua contribuição para aprimorar o desempenho dos detectores de emoção sem sensores (E2). Mais especificamente, o primeiro objetivo (E1) era investigar a dinâmica afetiva dos alunos brasileiros e verificar se diferentes informações sobre personalidade, motivação, sexo, comportamentos e duração das emoções poderiam influenciar no modelo de dinâmica afetiva. Os resultados de E1 indicam que a dinâmica afetiva dos estudantes brasileiros difere do modelo teórico de dinâmica afetiva de [D’Mello and Graesser 2012], sendo mais similar a pesquisas mais recentes, que consideram estudantes dos Estados Unidos [Karumbaiah et al. 2021]. No entanto, foram observadas diferenças significativas na maneira como os alunos transitam entre as emoções ao aprender com um STI. Em E1, também se explorou a influência de informações adicionais dos alunos nos modelos de dinâmica afetiva. Os resultados evidenciam que características adicionais dos alunos, como personalidade, motivação, sexo, comportamentos e informações sobre a duração das emoções, podem influenciar quais emoções os alunos experimentam e como elas são experimentadas ao aprender com um STI. Esses resultados sugerem que o modelo de dinâmica afetiva pode ser influenciado pelas informações pessoais dos alunos e pela duração das emoções.

O segundo objetivo (E2) era investigar se as informações sobre personalidade, motivação, sexo, comportamentos e dinâmica de afetos poderiam melhorar o desempenho dos detectores de emoção livres de sensores. Foi postulado que a inclusão dessas informações como características adicionais durante o desenvolvimento dos detectores de emoção resultaria em uma melhora do desempenho dos mesmos. Para testar essa hipótese, inicialmente foram desenvolvidos os modelos de referência para cada emoção, que consideraram apenas informações sobre a interação entre o aluno e o sistema (logs). Depois, foram desenvolvidos os modelos hipotéticos, que incorporaram, além dos logs, as informações adicionais dos alunos para cada emoção. Após a comparação do desempenho dos modelos de referência e hipotéticos utilizando o teste estatístico de McNemar, constatou-se que as informações sobre personalidade, motivação e sexo não melhoraram o desempenho na detecção de nenhuma emoção. Contudo, foi estatisticamente validado que informações sobre comportamentos do aluno, dados de dinâmica afetiva e a combinação de múltiplas fontes de dados aprimoraram o desempenho dos detectores de emoção.

Os resultados obtidos em ambos os estudos ressaltam que dados sobre personalidade, motivação, sexo e comportamentos do aluno não apenas influenciam a maneira como os alunos transitam entre as emoções (dinâmica afetiva), mas também podem melhorar o desempenho dos detectores de emoção livres de sensores. Assim, essas informações precisam ser consideradas por ambientes de aprendizagem sensíveis ao afeto. Ainda, esses resultados indicam que essas variáveis merecem ser investigadas com maior profundidade para entender melhor seu impacto no processo de aprendizagem.

## Agradecimentos

Este trabalho recebeu apoio financeiro das seguintes agências de fomento a pesquisa: CAPES (Código de Financiamento 001), FAPERGS (Processo 17/2551-0001203-8) e CNPq (processo 306005/2020-4).

## Referências

- Andres, J. M. A. L., Ocumpaugh, J., Baker, R. S., Slater, S., Paquette, L., Jiang, Y., Karumbaiah, S., Bosch, N., Munshi, A., Moore, A., et al. (2019). Affect sequences and learning in betty's brain. In *Proceedings of the 9th LAK*, pages 383–390.
- Arroyo, I., Woolf, B. P., Burelson, W., Muldner, K., Rai, D., and Tai, M. (2014). A multimedia adaptive tutoring system for mathematics that addresses cognition, meta-cognition and affect. *IJAIED*, 24(4):387–426.
- Azevedo, O., Morais, F., and Jaques, P. A. (2018). Exploring gamification to prevent gaming the system and help refusal in tutoring systems. In *European Conference on Technology Enhanced Learning*, pages 231–244. Springer.
- Baker, R. et al. (2010a). Data mining for education. *International encyclopedia of education*, 7(3):112–118.
- Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., and Graesser, A. C. (2010b). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, 68(4):223–241.
- Baker, R. S., Gowda, S. M., Wixon, M., Kalka, J., Wagner, A. Z., Salvi, A., Aleven, V., Kusbit, G. W., Ocumpaugh, J., and Rossi, L. (2012). Towards sensor-free affect detection in cognitive tutor algebra. *EDM*.
- Baker, R. S., Moore, G. R., Wagner, A. Z., Kalka, J., Salvi, A., Karabinos, M., Ashe, C. A., and Yaron, D. (2011a). The dynamics between student affect and behavior occurring outside of educational software. In *International Conference on Affective Computing and Intelligent Interaction*, pages 14–24. Springer.
- Baker, R. S., Moore, G. R., Wagner, A. Z., Kalka, J., Salvi, A., Karabinos, M., Ashe, C. A., and Yaron, D. (2011b). The dynamics between student affect and behavior occurring outside of educational software. In *International Conference on Affective Computing and Intelligent Interaction*, pages 14–24. Springer.
- Barbosa, A. A. G. (2009). *Modelo hierárquico de fobias infanto-juvenis: testagem e relação com os estilos maternos*. PhD thesis, Psicologia Social (UFRN).
- Bosch, N. and D'Mello, S. (2017). The affective experience of novice computer programmers. *IJAIED*, 27(1):181–206.
- Botelho, A. F., Baker, R. S., and Heffernan, N. T. (2017). Improving sensor-free affect detection using deep learning. In *AIED*, pages 40–51. Springer.
- Calvo, R. A. and D'Mello, S. K. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, 1(1):18–37.
- Conati, C. (2011). Combining cognitive appraisal and sensors for affect detection in a framework for modeling user affect. In *New perspectives on affect and learning technologies*, pages 71–84. Springer.
- de Araújo, D. C. S. A., Pereira, S. N., Dos Santos, W. M., dos Santos Menezes, P. W., Rocha, K. S. d. S., Cerqueira-Santos, S., Faro, A., Mesquita, A. R., and de Lyra Jr,

- D. P. (2021). Brazilian version of the personal report of communication apprehension: Cross-cultural adaptation and psychometric evaluation among healthcare students. *PLoS one*, 16(2):e0246075.
- D’Mello, S., Graesser, A., and Taylor, R. S. (2007). Monitoring affective trajectories during complex learning. In *Proceedings of the annual meeting of the cognitive science society*, volume 29.
- Dweck, C. S. and Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological review*, 95(2):256.
- D’Mello, S. and Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, 22(2):145–157.
- D’Mello, S., Lehman, B., Pekrun, R., and Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, 29:153 – 170.
- D’Mello, S. K., Craig, S. D., Witherspoon, A., Mcdaniel, B., and Graesser, A. (2008). Automatic detection of learner’s affect from conversational cues. *User modeling and user-adapted interaction*, 18(1-2):45–80.
- Fredrickson, B. L. (1998). What good are positive emotions? *Review of General Psychology*, 2(3):300–319.
- Frenzel, A. C., Pekrun, R., and Goetz, T. (2007). Girls and mathematics—a “hopeless” issue? a control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education*, 22(4):497–514.
- Goldoni, D. D., Reis, H. M., and Jaques, P. A. (2022). Modelagem estatística do tempo de permanência de estudantes no estado de confusão através de análise de sobrevivência multivariada. In *Anais do XXXIII Simpósio Brasileiro de Informática na Educação (SBIE 2022)*. Sociedade Brasileira de Computação - SBC.
- Gouveia, V. V., Diniz, P. K., Santos, W. S. d., Gouveia, R. S., and Cavalcanti, J. P. N. (2008). Metas de realização entre estudantes do ensino médio: evidências de validade fatorial e consistência interna de uma medida. *Psicologia: Teoria e Pesquisa*, 24(4):535–544.
- Graesser, A. and D’Mello, S. K. (2011). Theoretical perspectives on affect and deep learning. In *New perspectives on affect and learning technologies*, pages 11–21. Springer.
- Harley, J. M. (2016). Measuring emotions: a survey of cutting edge methodologies used in computer-based learning environment research. In *Emotions, technology, design, and learning*, pages 89–114. Elsevier.
- Henderson, N., Rowe, J., Paquette, L., Baker, R. S., and Lester, J. (2020). Improving affect detection in game-based learning with multimodal data fusion. In *International Conference on Artificial Intelligence in Education*, pages 228–239. Springer.
- Hutt, S., Grafsgaard, J. F., and D’Mello, S. K. (2019). Time to scale: Generalizable affect detection for tens of thousands of students across an entire school year. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–14.
- Jaques, P. A., Seffrin, H., Rubi, G., Morais, F., Ghilardi, C., Bittencourt, I. I., and Isotani, S. (2013). Rule-based expert systems to support step-by-step guidance in algebraic

- problem solving: The case of the tutor PAT2math. *Expert Systems with Applications*, 40(14):5456–5465.
- Jiang, Y., Bosch, N., Baker, R. S., Paquette, L., Ocumpaugh, J., Andres, J. M. A. L., Moore, A. L., and Biswas, G. (2018). Expert feature-engineering vs. deep neural networks: which is better for sensor-free affect detection? In *AIED*, pages 198–211. Springer.
- John, O. P. and Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, 2(1999):102–138.
- Karumbaiah, S., Baker, R. B., Ocumpaugh, J., and Andres, A. (2021). A re-analysis and synthesis of data on affect dynamics in learning. *IEEE Transactions on Affective Computing*.
- Linnenbrink-Garcia, L. and Barger, M. M. (2014). Achievement goals and emotions. *International handbook of emotions in education*, pages 142–161.
- Morais, F. and Jaques, P. A. (2021). Dinâmica de afetos dos alunos em um sistema tutor inteligente de matemática no contexto brasileiro. In *Anais do XXXII Simpósio Brasileiro de Informática na Educação*, pages 691–704. SBC.
- Morais, F. and Jaques, P. A. (2022). Dinâmica de afetos em um sistema tutor inteligente de matemática no contexto brasileiro: uma análise da transição de emoções acadêmicas. *Revista Brasileira de Informática na Educação*, 30:519–541.
- Morais, F. and Jaques, P. A. (2023). The dynamics of brazilian students’ emotions in digital learning systems (accepted for publication). *International Journal of Artificial Intelligence in Education*.
- Morais, F., Kautzmann, T. R., Bittencourt, I. I., and Jaques, P. A. (2019). Emap-ml: A protocol of emotions and behaviors annotation for machine learning labels. In *EC-TEL*, pages 495–509, Netherlands. Springer.
- Oakley, A. (2016). *Sex, gender and society*. Routledge.
- Ocumpaugh, J. (2015). Baker rodrigo ocumpaugh monitoring protocol (bromp) 2.0 technical and training manual. *New York, NY and Manila, Philippines: Teachers College, Columbia University and Ateneo Laboratory for the Learning Sciences*.
- Ocumpaugh, J., Baker, R., Gowda, S., Heffernan, N., and Heffernan, C. (2014). Population validity for educational data mining models: A case study in affect detection. *British Journal of Educational Technology*, 45(3):487–501.
- Paquette, L., Baker, R. S., Sao Pedro, M. A., Gobert, J. D., Rossi, L., Nakama, A., and Kauffman-Rogoff, Z. (2014). Sensor-free affect detection for a simulation-based science inquiry learning environment. In *ITS*, pages 1–10. Springer.
- Paquette, L., Rowe, J., Baker, R., Mott, B., Lester, J., DeFalco, J., Brawner, K., Sottolare, R., and Georgoulas, V. (2016). Sensor-free or sensor-full: A comparison of data modalities in multi-channel affect detection. In *EDM*. ERIC.
- Pardos, Z. A., Baker, R. S., San Pedro, M. O., Gowda, S. M., and Gowda, S. M. (2014). Affective states and state tests: Investigating how affect and engagement during the school year predict end-of-year learning outcomes. *JLA*, 1(1):107–128.

- Pekrun, R. (2014). Emotions and learning. In *Educational practices series*. IEA, IBE.
- Pekrun, R. (2016). Academic emotions. *Handbook of motivation at school*, 2:120–144.
- Rodrigo, M. M. T., d Baker, R. S., D’Mello, S., Gonzalez, M. C. T., Lagud, M. C., Lim, S. A., Macapanpan, A. F., Pascua, S. A., Santillano, J. Q., Sugay, J. O., et al. (2008). Comparing learners’ affect while using an intelligent tutoring system and a simulation problem solving game. In *International Conference on Intelligent Tutoring Systems*, pages 40–49. Springer.
- Sabourin, J., Rowe, J. P., Mott, B. W., and Lester, J. C. (2011). When off-task is on-task: The affective role of off-task behavior in narrative-centered learning environments. In *International Conference on Artificial Intelligence in Education*, pages 534–536. Springer.
- Sinclair, J., Jang, E. E., Azevedo, R., Lau, C., Taub, M., and Mudrick, N. V. (2018). Changes in emotion and their relationship with learning gains in the context of metatutor. In *ITS*, pages 202–211. Springer.
- Tong, E. M. (2010). Personality influences in appraisal–emotion relationships: The role of neuroticism. *Journal of Personality*, 78(2):393–417.
- Wolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D., and Picard, R. (2009). Affect-aware tutors: Recognising and responding to student affect. *International Journal of Learning Technology*, 4(3/4):129–164.