CBIE 2023
CONGRESSO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO
Uma escola para o futuro: Tecnologia e conectividade a serviço da educação

Do gamified tutoring systems hinder sexual diversity? An experimental study with cis-heteronormative stereotype

Breno Felix de Sousa[1], Geiser Chalco Challco[2], Marcelo Reis[1], Ig Ibert Bittencourt Santana[1], Jessica Fernanda Silva Barbosa[1], Francys Rafael do Nascimento Martins[1], Seiji Isotani[2]

[1]Instituto de Computação – Universidade Federal de Alagoas (UFAL)
Caixa Postal S/N – 57.072.-900 – Maceió – AL – Brazil

[2]Instituto de Ciências Matemáticas e de Computação – Universidade de São Paulo (USP)
São Carlos, Brazil.

{bfs, jfsb, frnm, ig.ibert,}@ic.ufal.br, geiser@alumni.usp.br

sisotani@icmc.usp.br, marceloreis.bio@gmail.com

Abstract. *Cis-heteronormative stereotyping in STEM fields is responsible for making minority groups like LGBTQ+ feel like they don't belong in these fields of science. To identify and analyze the effects of such stereotypes, we performed a quasi-experiment with N = 70 participants in a gamified system. The results indicate that self-declared cis-heterosexual participants had better flow experiences in environments with stereotypes that disagreed with their sexual orientation. As for the expected performance, stereotypes not aligned with sexual orientation were harmful. These results reflect the importance of developing socially inclusive educational technologies, with mechanisms that allow customization of the interface to avoid disadvantage of minority groups.*

1. Introduction

Cis-heteronormativity is the dominant status in the fields of Science. Technology, Engineering, and Mathematics (STEM). These fields perpetuate the notion that anything that goes against this hegemonic status does not belong [Powell et al. 2020]. STEM is a masculine culture [Miller et al. 2021]. To comprehend the obstacles faced by minority groups such as LGBTQ+ (Lesbian, Gay, Bisexual, Transgender, and Queer or Questioning) in these scenarios, we must examine how oppressive standards are rooted in historical anti-LGBTQ+ policies [Alexander et al. 2022]. This implies the need to recognize that LGBTQ+ struggles and history are often disregarded in the scientific community [Maguth and Taylor 2014].

With the COVID-19 pandemic, the popularity of virtual education has increased and promoted the interest in developing efficient tools for online learning [Dwivedi et al. 2020]. New educational technology must seek to promote sexual diversity and avoid perpetuating cis-heteronormative stereotypes, especially in the fields of STEM.

Scientific studies have shown that including and amplifying queer voices with interactive technology design can lead to transforming reflections that can benefit game design and gamification [Hantsbarger et al. 2022]. Replicating traditional teaching patterns to virtual learning settings may contribute to excluding minority LGBTQ+ groups

in predominantly cis-heteronormative spaces. This concern becomes more relevant when considering virtual learning systems' development and gamification. Gamification is defined as using game design elements in non-gaming scenarios and is an approach that can motivate and improve students' learning results [Klock 2020]. In these platforms, cis-heteronormative stereotypes may be introduced. This may occur when the platform's design elements, such as exclusively male or female avatars, do not resonate with the LGBTQ+ participants' sexual orientations.

This study aimed to verify whether sexual stereotypes (cis-heteronormative and LGBTQ+ stereotypes) in gamified tutoring systems affect expected performance. The effects of stereotypes on flow state were also investigated. Flow state is a mental disposition sought by educators and educational technologies and consists of a state of complete immersion in a task [Alameda et al. 2022]. In this sense, it is relevant to understand the effects of stereotypes in gamified platforms on expected learning performances [Shaw-Zirt et al. 2005], considering that such stereotypes have a negative impact on students [Monterrat et al. 2017].

2. Related studies

Gamification and stereotypes. Stereotypes introduced in gamified systems may influence participants' performances. Therefore, this aspect should be considered while developing gamified learning platforms. An investigation on gender stereotype threat (male and female stereotypes) in gamified platforms has been previously conducted by Albuquerque et al. [Albuquerque et al. 2017]. In the study, anxiety levels varied according to gender and stereotypes. Changes in anxiety were higher among females who were under male stereotype threat.

Albuquerque et al.[Albuquerque et al. 2017] did not delve into sexual diversity and approached only anxiety and flow state levels. In our study, we intend to investigate the effects of another type of stereotype: sexual stereotypes. That is cis-heteronormative stereotypes (e.g., cis-heteronormative avatars) and LGBTQ+ stereotypes (e.g., lesbian, gay, bisexual, transgender avatars). We also expanded the scope of the previous study by analyzing the effects of stereotypes on expected performance.

Expected performance is not necessarily related to the performance itself, as it may be understood as the intensity of someone's belief in their ability to perform a task [Yu et al. 2021]. Still, a more intense belief in being able to accomplish the task successfully tends to lead to better performance [Taiwo et al. 2012].

In a previous experimental study, gender stereotypes had an important influence on performance [Khusaini and Mulya 2021]. The performance variables confirmed that female studies did not achieve excellent results when facing male gender stereotypes. With this, it was suggested that stereotypes could influence performance.

Sexual stereotypes are blatantly present in STEM settings. Consequently, minority groups not aligned with the dominant status tend to feel that they do not belong in these scenarios [Shaw-Zirt et al. 2005]. Evasion and low performance are also some of the repercussions [Hughes and Kothari 2021]. Exercising perception may help to understand that sexual segregation exceeds traditional and stigmatized barriers within STEM fields [Powell et al. 2020]. The LGBTQ+ students who participated in this study reported their

struggles to fight prejudice, how colleagues can be allies in this fight, and what policies the institutions should implement to make STEM settings more inclusive.

A study that used data from 21 professional societies (with 25,324 cisgender people and 1006 LGBTQ+ people) showed that LGBTQ+ professionals are more likely to have career limitations and be harassed and devalued [Cech and Waidzunas 2011]. Consequently, these groups are reported to be more prone to dropping out of these programs. What can potentially explain these results is that stereotype threat can decrease performance expectations, which could, in turn, lead to worse performances [Cadinu et al. 2003].

3. Method

This quasi-experiment was performed to answer the research question: "do sexual stereotypes (cis-heteronormative and LGBTQ+ stereotypes) impact flow experience, expected performance, and performance in gamified tutoring systems?" With this intent, we formulated the following null hypotheses:

- H1: There is no significant difference in the flow state of participants according to their sexual orientation (cis-heterosexual and LGBTQ+) and environments with cis-heteronormative stereotypes (not.LGBTQ+) and LGBTQ+ stereotypes (stLGBTQ+);
- H2: There is no significant difference in the expected performance of participants according to their sexual orientation (cis-heterosexual and LGBT+) and environments with cis-heteronormative stereotypes (not.LGBTQ+) and LGBTQ+ stereotypes (stLGBTQ+);
- H3:There is no significant difference in the performance of participants according to their sexual orientation (cis-heterosexual and LGBT+) and environments with cis-heteronormative stereotypes (not.LGBTQ+) and LGBTQ+ stereotypes (stLGBTQ+).

4. Research Design

The study used a 2x2 factorial design with two conditions and two factors per condition. The factors are a gamified interface with LGBTQ+ stereotypes (stLGBTQ+) and a gamified interface with cis-heteronormative stereotypes (non.LGBTQ+). The participants were assigned to one of two conditions: the LGBTQ+ intervention group or the control heterosexual group. The LGBTQ+ and heterosexual conditions were associated with the type of interface factor: stLBTQ+ (intervention) and non.LGBTQ+ (control). The gamified tutoring system randomly assigned each participant to one of two factors: LGBTQ+ stereotype (stLGBTQ+) or heterosexual (non.LGBTQ+). In figure 1, the experiment's execution flow is seen. We used a web-based system that required only a link to access. After receiving the link to the experiment, the participant was asked to sign the free and informed consent term (TCLE), which the ethics committee requires for research with human beings. By agreeing to participate in the research, the participant was redirected to the pre-test phase, which consisted of answering the DFS-Short BR (dfs) and the sexual diversity questionnaires. Then, each participant was randomly redirected to one of the two interfaces (stLGBTQ+ or non.LGBTQ+). At this point, the exposure to stereotyped gamified elements began. The user was required to choose an avatar to move forward to

the quiz, which was composed of 20 logic/mathematics questions. By finishing the quiz, the participant was directed to the post-test phase, which consisted of answering the FSS-2 (fss) and the socioeconomic questionnaire. No time limit was set for the completion of the questionnaires.

The control interface (non.LGBTQ+) had elements with heterosexual stereotypes, such as cis-heteronormative avatars. While the intervention interface (stLGBTQ+) was composed of LGBTQ+ stereotypes, such as avatars associated with sexual diversity. The control setting also presented stereotyped boost phrases, colors, sound effects, and bar graphs to convey the message that heterosexuals have better performances in logic/mathematics activities. As for the intervention setting, the stereotypes conveyed the feeling of sexual diversity and affirmed that LGBTQ+ groups perform better in logic/mathematics activities. The choice of these stereotypes was intentional and based on the literature Male-stereotyped posters and figures, such as Star Wars or Star Trek, and stereotyped colors, such as blue and gray, are traditional within STEM environments [Ferraz and Gama 2019]. Considering this, the control interface was designed based on social assumptions and perceptions prevalent within the STEM fields. As for the LGBTQ+ interface, it was conceived to include a broad array of sexual stereotypes.

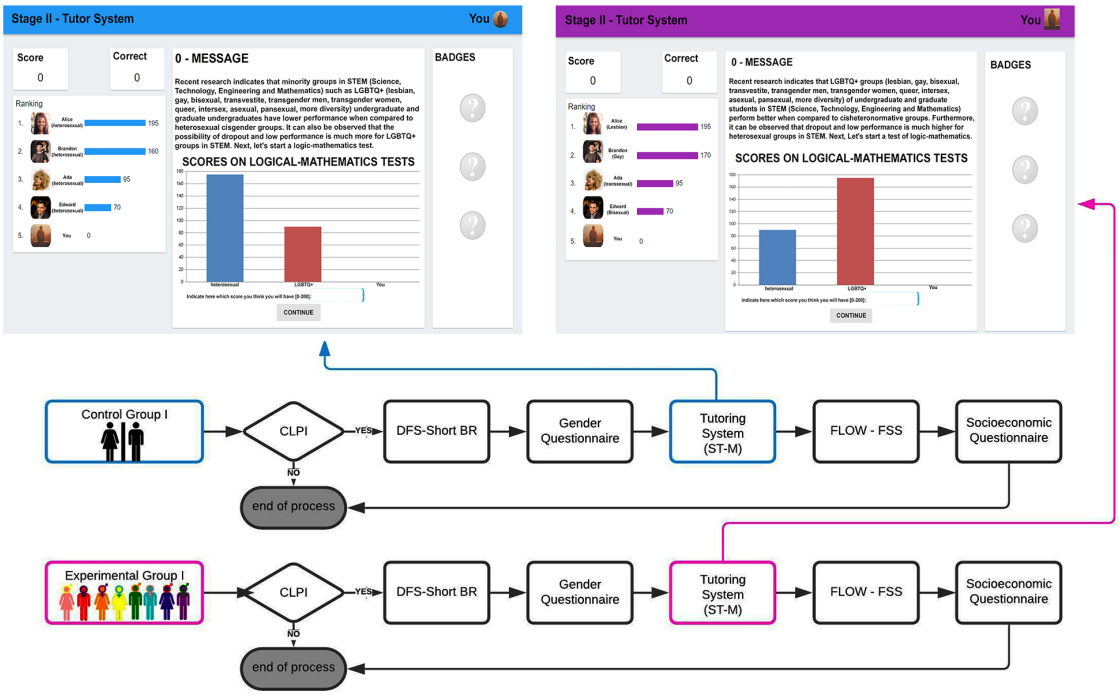

Figura 1. Study design and data collection procedures

4.1. Participants and context of the study

The participants were 70 undergraduate and postgraduate students from STEM fields at the Federal University of Alagoas (UFAL) in the Maceió campus in Alagoas, Brazil. Regarding gender, 74.28% (n = 52) of participants were men, and 25.72% (n=18) were women. Concerning sexual orientation, 80.28% (n = 57) were heterosexual, 5.64% (n = 4) were homosexual, 11.26% (n = 8) were bisexual, 1.41% (n = 1) were asexual, and 1.41%

(n= 1) were pansexual. As for gender identity, 71.84% (n = 51) were cisgender men, 18.31% (n = 13) were cisgender women, and 9.85% (n = 7) did not answer. In relation to ethnicity, 45.07% (n = 32) of participants self-declared as Brown, 38.02% (n = 27) self-declared as White, 11.26% (n = 8) self-declared as Black, and 5.65% (n = 4) chose not to answer. Brown is a term used by the Brazilian Institute of Geography and Statistics (IBGE) that refers to one of the five racial groups that make up the Brazilian population: White, Black, Asian, Indigenous, and Brown [de Geografia e Estatística 2020].

4.2. Data Collection and Analysis

We used the translated and validated Brazilian Portuguese versions of the DFSBR and FSS-BR. The version translated to Brazilian Portuguese uses a five-point Likert scale, ranging from 1 (never) to 5 (always). Developed and validated a nine-question short version of the original DFS-2, translated into Brazilian Portuguese [Bittencourt et al. 2021] [Jackson et al. 2008]. In this study, we used the DFS-Short-BR version to optimize our experiment's execution.

The degree of expected performance was based on the work of [Cadinu et al. 2003]. The expected performance assessment consisted of asking the student a previous estimation of how they would perform in the test, which could range from 0 to 200. The tutoring system was controlled to ensure participants were assigned to the interfaces according to their sexual orientations. The goal was to predict whether negative or positive information could affect expected performance. The performance estimations made by participants were used to measure the expected performance deficit.

Learning performance was measured with the number of points obtained using the platform. For each correct answer, 10 points were added to the participant's score. This scoring system allowed for checking participants' performances in the experiment and comparing the two types of interfaces.

The participants had no knowledge of the existence of two types of settings. Thus, the study consisted of a single-blinded experiment and participation was voluntary. Due to the lack of knowledge as to who was performing the task, we also expected the absence of interactions between students within the two different settings. The experiment was approved by the human research ethics committee.

5. Findings and Discussion

Table 1 presents the descriptive statistics and estimated marginal means (adjusted) for dispositional flow (dfs) and flow scale (fss), for expected performance and performances grouped according to sexual orientation (heterosexual versus LGBTQ+), and for setting type. The control environment was the non.LGBTQ+ (with heterosexual stereotypes), while the intervention setting was the stLGBTQ+ (with LGBTQ+ stereotypes).

Table 2 ndicates the ANCOVA test results for flow experience and the ANOVA test results for expected performance and performance. The tests were conducted to assess whether there was a significant difference between participants in the stLGBTQ+ setting and participants in the non.LGBTQ+ one.

Tabela 1. Descriptive statistics of student flow status, expected performance, and learning performance based on sexual orientation

		DFS (before)			FSS (after)		Expected performance			Activity Points		
stType	**Orientation**	**N**	**M**	**SE**	**M**	**SE**	**N**	**M**	**SE**	**N**	**M**	**SE**
stLGBTQ+	LGBTQ+	9	3.452	0.192 3	3.656	0.244	9	118.444	9.149	9	127.222	17.836
	non.LGBTQ+	22	3.518	0.124	4.161	0.112	22	102.414	5.852	22	152.955	4.170
stHetero	LGBTQ+	8	3.764	0.142	4.107	0.159	8	123.719	9.704	8	118.125	16.901
	non.LGBTQ+	31	3.466	0.134	3.754	0.106	31	123.258	4.930	31	133.387	6.985

Tabela 2. Descriptive statistics of student flow state, expected performance, and learning performance based on environment stereotypes setting

	DFS (before)			FSS (after)		Expected performance			Activity Points		
Condition	**N**	**M**	**SE**	**M**	**SE**	**N**	**M**	**SE**	**N**	**M**	**SE**
stThreat	30	3.566	0.102	4.077	0.099	21	107.810	6.183	30	143.667	6.990
stBoost	40	3.464	0.088	3.732	0.101	39	124.821	4.537	40	132.000	6.054

5.1. H1: Flow experience per sexual orientation and gamified settings

After controlling the linearity of the ”dfs”covariate, we performed ANCOVA tests with the independent between-subjects ”setting”(non.LGBTQ+ and stLGBTQ+) and ”group”(LGBTQ+ and heterosexual) variables to determine if there were statistically significant differences in the dependent ”fss”variable. For the dependent ”fss”variable, there were statistically significant effects in the ”dfs”factor (F(1.63)=11.318, p=0.001, ges=0.152 [effect size]) and in the interaction between factors ”setting:group”(F(1.63)=5.898, p=0.018, ges=0.086 [effect size]).

Paired comparisons using Estimated Marginal Means (EMMs) were computed to find statistically significant differences between groups defined by the independent variables. The p-values were adjusted through the ”bonferroni”method. For the dependent variable ”fss”, the mean for the setting=”not.LGBTQ+”(adj M=3.771 and SD=0.589) was significantly different compared to the mean for setting=“stLGBTQ+” (adj M=4.16; SD=0.502; p-adj=0.013). The mean for group=“LGBTQ+” (adj M=3.677 and SD=0.731) was significantly different compared to the mean for group=“not.LGBTQ+” (adj M=4.16; SD=0.502; p-adj=0.027).

This indicates that students had higher flow levels in an environment that disagreed with their sexual orientation. It is possible that when under stereotype threat, participants refuted the stereotyped setting, resulting in a higher immersion in the system. However, self-declared heterosexuals had a higher flow variation when under threat compared to LGBTQ+ participants. Based on the literature, we assume this may reflect the ”global closet”[Pachankis and Bränström 2019]. When in a stereotypethreat environment, self-declared heterosexual participants could, in fact, have been in a setting that corresponded to their sexual orientation, despite having concealed it. The similar variations in flow state within the non.LGBTQ+ among heterosexuals and LGBTQ+ participants obtained in the study support this assumption. In the stLGBTQ+ setting, the variation in flow levels was higher for self-declared heterosexuals. Thus, null hypothesis H1 was rejected.

5.2. H2: Expected performance per sexual orientation and gamified setting

We conducted ANOVA tests with the independent between-subjects variables ”setting”(non.LGBTQ+ and stLGBTQ+) and ”group”(LGBTQ+ and heterosexual) to deter-

mine whether there were statistically significant differences in the dependent ”expected performance”variable. For the dependent ”expected performance”variable, we found statistically significant effects for the ”setting”factor (F(1,66)=6,552; p=0,013; ges=0,09 [effect size]).

Paired comparisons using Estimated Marginal Means (EMMs) were computed to find statistically significant differences between groups defined by the independent variables. The p-values were adjusted through the ”bonferroni”method. For the dependent “expected performance” variable, the mean for setting= “not.LGBTQ+” (adj M=123.258 and SD=30.362) was significantly different from the mean for setting= “stLGBTQ+” (adj M=102.414; SD=23.032; p-adj=0.008).

Hence, we concluded that the null hypothesis H2 was rejected. The LGBTQ+ and the heterosexual groups, when in a setting aligned with their sexual orientation, had higher expected performance rates. However, when in an environment in disagreement with their sexual orientation, both groups obtained lower expected performance rates. Nevertheless, the variations in mean expected performance for LGBTQ+ groups were lower in both conditions (boost and threat), but were higher in a setting in agreement with their sexual orientation, which reinforces the importance of inclusive gamified environments.

5.3. H3: Learning performance per sexual orientation and gamified setting

The ANOVA tests with the independent between-subjects ”setting”(not.LGBTQ+ e stLGBTQ+) and ”group”(LGBTQ+ and heterosexual) variables were performed to determine whether there were statistically significant differences for the dependent ”points”variable. There were no statistically significant effects for the dependent ”points”variable. These test results indicate no statistically significant effects for the ”orientation”factor (F(1,66)=6,932; p=0,056; ges=0,105 [effect size]).

Thus, null hypothesis H3 was accepted. Therefore, we concluded that there were no significant differences in learning performance according to sexual orientation. The literature proposes that specific characteristics of a virtual setting can negatively affect minority groups’ psychological mediators. Such effects do not always impact learning results but could become an obstacle for these groups [Powell et al. 2020].

6. Conclusions

The experiment demonstrated that sexual stereotypes affect flow experience and expected performance levels. However, for both the LGBTQ+ and the heterosexual groups, learning performance was not affected. It was seen that, regardless of sexual orientation, participants under threat presented better flow experiences. This study instigates further empirical and qualitative investigations to promote a better understanding of the observed phenomena. In a practical scenario, the study reinforces the need for developing gamified interfaces that consider sexual stereotypes to ensure optimal expected performances for users. Because the STEM fields are highly associated with cis-heteronormative stereotypes, the design of gamified learning technologies should avoid replicating these traditional stigmas

7. References

Referências

Alameda, C., Sanabria, D., and Ciria, L. F. (2022). The brain in flow: a systematic review on the neural basis of the flow state. *Cortex*.

Albuquerque, J., Bittencourt, I. I., Coelho, J. A., and Silva, A. P. (2017). Does gender stereotype threat in gamified educational environments cause anxiety? an experimental study. *Computers & Education*, 115:161–170.

Alexander, N., Knutson, D., Lynch, L., Spellman, M., Rivera, M., Morrow, L. K., Scherz, M., Clarkberg, T., Allen-Custodio, L. T., and Fountain, K. (2022). Increasing inclusion & competency in stem: Understanding lgbtq+ history, barriers, and heteronormativity.

Bittencourt, I. I., Freires, L., Lu, Y., Challco, G. C., Fernandes, S., Coelho, J., Costa, J., Pian, Y., Marinho, A., and Isotani, S. (2021). Validation and psychometric properties of the brazilian-portuguese dispositional flow scale 2 (dfs-br). *PloS one*, 16(7):e0253044.

Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., and Latinotti, S. (2003). Stereotype threat: The effect of expectancy on performance. *European journal of social psychology*, 33(2):267–285.

Cech, E. A. and Waidzunas, T. J. (2011). Navigating the heteronormativity of engineering: The experiences of lesbian, gay, and bisexual students. *Engineering Studies*, 3(1):1–24.

de Geografia e Estatística, I. B. (2020). Pnad contínua - pesquisa nacional por amostra de domicílios contínua.

Dwivedi, Y. K., Hughes, D. L., Coombs, C., Constantiou, I., Duan, Y., Edwards, J. S., Gupta, B., Lal, B., Misra, S., Prashant, P., et al. (2020). Impact of covid-19 pandemic on information management research and practice: Transforming education, work and life. *International journal of information management*, 55:102211.

Ferraz, C. and Gama, K. (2019). A case study about gender issues in a game jam. In *Proceedings of the International Conference on Game Jams, Hackathons and Game Creation Events 2019*, pages 1–8.

Hantsbarger, M., Troiano, G. M., To, A., and Harteveld, C. (2022). Alienated serendipity and reflective failure: Exploring queer game mechanics and queerness in games via queer temporality. *Proceedings of the ACM on Human-Computer Interaction*, 6(CHI PLAY):1–27.

Hughes, B. E. and Kothari, S. (2021). Don't be too political: depoliticization, sexual orientation, and undergraduate stem major persistence. *Journal of homosexuality*, pages 1–28.

Jackson, S. A., Martin, A. J., and Eklund, R. C. (2008). Long and short measures of flow: The construct validity of the fss-2, dfs-2, and new brief counterparts. *Journal of Sport and Exercise Psychology*, 30(5):561–587.

Khusaini, K. and Mulya, A. S. (2021). Improving the students'performance from the expectation and quality of learning. *Assets: Jurnal Akuntansi dan Pendidikan*, 10(2):147–164.

Klock, G. (2020). Pimenta and hamari, 2020 klock act, gasparini i., pimenta ms, hamari j. *Tailored gamification: A review of literature, International Journal of Human-Computer Studies*, 144:102495.

Maguth, B. M. and Taylor, N. (2014). Bringing lgbtq topics into the social studies classroom. *The Social Studies*, 105(1):23–28.

Miller, R. A., Vaccaro, A., Kimball, E. W., and Forester, R. (2021). "it's dude culture": Students with minoritized identities of sexuality and/or gender navigating stem majors. *Journal of Diversity in Higher Education*, 14(3):340.

Monterrat, B., Lavoué, É., and George, S. (2017). Adaptation of gaming features for motivating learners. *Simulation & Gaming*, 48(5):625–656.

Pachankis, J. E. and Bränström, R. (2019). How many sexual minorities are hidden? projecting the size of the global closet with implications for policy and public health. *PLoS One*, 14(6):e0218084.

Powell, K., Terry, R., and Chen, S. (2020). How lgbt+ scientists would like to be included and welcomed in stem workplaces. *Nature*, 586(7831):813–817.

Shaw-Zirt, B., Popali-Lehane, L., Chaplin, W., and Bergman, A. (2005). Adjustment, social skills, and self-esteem in college students with symptoms of adhd. *Journal of attention disorders*, 8(3):109–120.

Taiwo, A. A., Mahmood, A. K., and Downe, A. G. (2012). User acceptance of egovernment: Integrating risk and trust dimensions with utaut model. In *2012 international conference on computer & information science (ICCIS)*, volume 1, pages 109–113. IEEE.

Yu, L., Chen, Z., Yao, P., and Liu, H. (2021). A study on the factors influencing users' online knowledge paying-behavior based on the utaut model. *Journal of theoretical and applied electronic commerce research*, 16(5):1768–1790.