Attractive : Innovative Education Journal
Vol. 8 No. 4, December 2026

Trust in Platform and Parental Control Adoption on Roblox in Indonesia: A Technology Threat Avoidance Theory Approach

Wisnu Wardana [1*], Ahmad Nurul Fajar [1]
[1] *Bina Nusantara University Jakarta, Indonesia*

wisnuyare@binus.ac.id*

ARTICLE INFO
Article history:
Received
August 12, 2026
Revised
October 13, 2026
Accepted
November 13, 2026

ABSTRACT

This study analyzes the factors that influence parental control adoption on Roblox by integrating Trust in Platform into Technology Threat Avoidance Theory (TTAT). It addresses the gap between the availability of parental control tools and their use by parents: only 38.5% of the surveyed parents reported having activated them. A cross-sectional survey of 257 parents in Jabodetabek was analyzed with PLS-SEM (SmartPLS 4.0; 5,000 bootstrap subsamples) and specific indirect effects. Five of seven hypotheses were supported. Trust in Platform positively influences Safeguard Effectiveness ($\beta = 0.441$) and Self-Efficacy ($\beta = 0.461$), and its indirect effect on Avoidance Motivation runs through Self-Efficacy rather than Safeguard Effectiveness. Perceived Threat is the strongest predictor of Avoidance Motivation ($\beta = 0.626$, $f^2 = 0.548$). Safeguard Effectiveness had no effect, and Safeguard Cost had a small positive effect, which we interpret as consistent with stakes-override and cost-as-signal dynamics. The study provides evidence that platform trust is a TTAT antecedent in a captive-market, proxy-protection setting, and derives feature-design implications (guided setup, age-based presets, protective defaults) and a three-pillar policy ecosystem (regulators, schools, developers) for PP No. 17/2025 and PM Komdigi No. 9/2026.

Keywords: Trust in Platform, Technology Threat Avoidance Theory, Parental Control, Roblox, Child Digital Safety

Journal Homepage https://attractivejournal.com/index.php/aj/
This is an open access article under the CC BY SA license
https://creativecommons.org/licenses/by-sa/4.0/
@ 2026 by the authors
Published by CV. Creative Tugu Pena

INTRODUCTION

Roblox has grown from a game into something closer to a social platform, where children build, chat and spend real money inside a space largely shaped by other users. Children under 13 alone logged around 9 billion hours on it in the second quarter of 2025 (Statista, 2025). That scale carries risk. Research on online games documents cyberbullying among adolescent players (Huang et al., 2019). A UNICEF baseline study of children aged 8–18 in three Indonesian provinces found that almost half had experienced some form of cyberbullying (UNICEF Indonesia, 2023), and *Disrupting Harm* estimated that at least 2% of internet-using Indonesian children aged 12–17 had been subjected to clear examples of online sexual exploitation and abuse in the year before its survey, a figure its authors consider likely to reflect underreporting (ECPAT et al., 2022). Studies of virtual worlds such as Roblox and Minecraft also find that their communication and parental-control features are often misaligned with how children actually socialize (Du et al., 2021).

Roblox offers a suite of parental controls, including account restrictions, contact settings, billing restrictions and playtime tracking. On paper, parents have what they need. In practice, the evidence points to a gap between availability and use. The same UNICEF study found that most of the 509 parents it interviewed felt they did not know enough about online safety to help their children, and that monitoring was mostly informal: checking the child's phone (56%), asking questions and advising (17.3%), or accompanying the child online (8.7%) (UNICEF

Indonesia, 2023). We found no national figure for how many Indonesian parents use platform-level parental controls, so this study measures the gap directly. Only 38.5% of our respondents reported having activated them on a single self-report item, even though every respondent met an inclusion criterion of basic awareness of the features. Abroad, a nationally representative survey in the United States found that use of parental controls ranged from 35% on game consoles to 51% on tablets, that more than half of parents did not use them on each device studied, and that more than four in five households using them found them effective (Family Online Safety Institute & Ipsos, 2025). The tools exist, and where they are used they are seen to work. Cyber parental control is an established research field (Altarturi et al., 2020), yet parents often see the tools as monitoring-oriented and are put off by usability and privacy and security gaps (Gnanasekaran & De Moor, 2025). Earlier work has explained parental intervention through mediation style, digital skills, self-efficacy and worry about online risks (Glatz et al., 2018; Livingstone et al., 2017; Livingstone & Helsper, 2008; Sciacca et al., 2022), and parental control apps have been studied as a design space that parents may read as either protection or punishment (Wang et al., 2021; Wisniewski et al., 2017). Work in Indonesia points to parental digital illiteracy and a communication paradox around children's video game practices (Fajar et al., 2025). What remains unexplained is why parents who are aware of a safeguard still do not turn it on.

Technology Threat Avoidance Theory (TTAT) offers a useful starting point. It explains avoidance behavior as the product of threat appraisal and coping appraisal (Liang & Xue, 2009, 2010), has been refined and tested in several settings (Carpenter et al., 2019; D. Q. Chen & Liang, 2019; Sreenath et al., 2025), and has been applied to at least one other Indonesian digital-threat setting, the avoidance of QR-code payment scams (Alpaullivarez & Fajar, 2026). TTAT was not built for a scenario like Roblox, however. A parent frustrated with Roblox cannot switch to a better safety tool the way a user might switch antivirus software. Device-level controls cannot see inside the game, so they cannot govern who messages a child, what appears in chat or how Robux is spent. Only Roblox's own systems reach that far. When the company that created the risk is also the only one able to mitigate it, a parent's willingness to act is likely to depend on whether they trust that company. TTAT, as originally formulated, has no place for that variable.

This study adds it. We integrate Trust in Platform (Gefen et al., 2003; McKnight et al., 2002) into TTAT and test the extended model with PLS-SEM on a sample of 257 parents in Jabodetabek. Trust has been studied in Indonesian technology adoption research (Arsyad & Fajar, 2026; Pertiwi et al., 2025; Sebayang et al., 2024), but rarely as an antecedent inside a threat-avoidance model, and, as far as we can tell, not in the context of child digital safety. The study has three aims: to test the extended model, to identify the mechanism through which trust operates, and to turn the findings into recommendations for regulators implementing PP No. 17/2025 (Pemerintah Republik Indonesia, 2025) and PM Komdigi No. 9/2026 (Kementerian Komunikasi dan Digital Republik Indonesia, 2026), and for platforms.

Theoretical background and hypotheses

In TTAT, perceived threat and coping appraisal jointly determine avoidance motivation, which in turn drives avoidance behavior (Liang & Xue, 2009, 2010). Coping appraisal has three components: the perceived effectiveness of the safeguard, the perceived cost of using it, and the user's self-efficacy. In this study the safeguard is Roblox's parental control. Trust in Platform is placed before the two coping beliefs that depend most on what the provider does, namely effectiveness and self-efficacy. The conceptual model is shown in Figure 1.

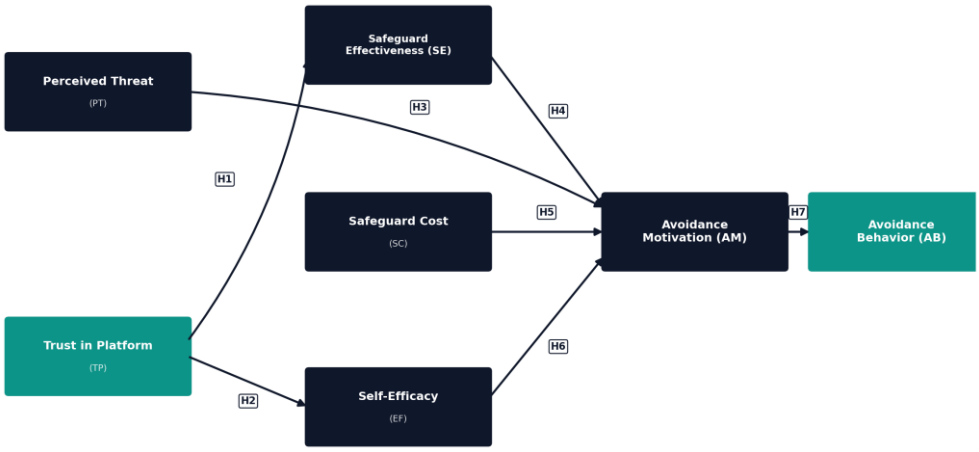
Trust and coping appraisal (H1, H2). Trusting beliefs rest on perceived competence, benevolence and integrity (McKnight et al., 2002), and trust is a known antecedent of beliefs about a technology's usefulness (Gefen et al., 2003). A parent who believes Roblox is competent and acts in users' interest should also believe that its safety features work (H1). Platform design matters for self-efficacy as well: among Chinese teenage gamers, easy-to-use platform privacy functions raised protective self-efficacy (H. Chen et al., 2024), and in Sri Lanka, parents' internet self-efficacy and trust in other online users shaped parental intervention (Tennakoon et al., 2018). We therefore expect trust to raise parents' confidence that they can use the controls (H2).

Threat and coping appraisal (H3 to H6). Perceived threat is the core driver of avoidance motivation in TTAT (Liang & Xue, 2009), and parents who perceive severe risks are more likely

to mediate their children's media use (Hwang et al., 2017). Perceived effectiveness should raise motivation, as response efficacy does in protection motivation studies (Hwang et al., 2017; Jansen & van Schaik, 2018). Safeguard cost should lower it, because cost is a barrier to anti-malware behaviors (Blythe & Coventry, 2018) and a negative predictor in security studies (Han et al., 2025). Self-efficacy (Bandura, 1977) should raise it (Glatz et al., 2018; Jansen & van Schaik, 2018).

Motivation and behavior (H7). Intention is a proximal predictor of security behavior (Anderson & Agarwal, 2010), although the intention-behavior gap in security can be wide, especially when the required effort is high (Jenkins et al., 2021). Based on this reasoning we test the following hypotheses: H1: Trust in Platform (TP) positively affects Safeguard Effectiveness (SE). H2: TP positively affects Self-Efficacy (EF). H3: Perceived Threat (PT) positively affects Avoidance Motivation (AM). H4: SE positively affects AM. H5: Safeguard Cost (SC) negatively affects AM. H6: EF positively affects AM. H7: AM positively affects Avoidance Behavior (AB).

Figure 1. Conceptual research model with hypotheses (H1–H7)



METHOD

This study used a quantitative, cross-sectional survey design (Creswell & Creswell, 2018). The target population was parents or guardians in Jabodetabek with children aged 7–17 who play Roblox at least once a week. Purposive sampling was applied with four inclusion criteria, one of which was basic parental awareness of Roblox's parental control features. This criterion introduces a selection effect toward parents who are aware of the tools but vary in whether they use them. Data collection yielded 257 valid responses between April and May 2026. Respondents were recruited through WhatsApp parenting groups, Twitter (with a Robux-based incentive), LinkedIn and Facebook, and through snowball distribution in personal networks.

The instrument had 28 items across seven constructs, each rated on a five-point Likert scale. Items for the TTAT constructs were adapted from Liang & Xue (2010), and items for Trust in Platform from the trusting-beliefs scale of McKnight et al. (2002), all tailored to Roblox. Appendix A shows representative items. The study did not require ethics committee approval under institutional policy for non-invasive, anonymous surveys of adults. Participation was voluntary, responses were anonymous, and each participant was told the purpose of the research before agreeing to take part.

Sample size was checked in two ways. First, an a priori power analysis targeted the most demanding path, the prediction of AM from four predictors (PT, SE, SC and EF). With $\alpha = .05$, power = .80 and four predictors, a medium effect ($f^2 = 0.15$) requires only 85 cases, and a noncentral-F calculation shows that N = 257 can detect effects of $f^2 \geq 0.047$. Second, the inverse square root method (Kock & Hadaya, 2018) indicates that N = 257 detects path coefficients of about 0.155 or larger at 80% power. Small effects below these thresholds, such as the H5 path, are therefore at the edge of what this sample can detect.

The data were analyzed in SmartPLS 4.0, chosen for the model's exploratory orientation, its moderate sample and its structural complexity (Hair et al., 2019, 2022). The outer model was evaluated with loadings ≥ 0.70, AVE ≥ 0.50, CA and CR ≥ 0.70, the Fornell-Larcker criterion and

HTMT < 0.85 (Henseler et al., 2015). The inner model was evaluated with bias-corrected bootstrapping (5,000 subsamples; $t > 1.96$, $p < .05$), R^2, Q^2, f^2 and inner-model collinearity (VIF < 5.0). Common method bias was assessed with full-collinearity VIF (Kock, 2015) and interpreted with the cautions in Podsakoff et al. (2024). Specific indirect effects were used to test mediation. Figure 2 summarizes the procedure.

Figure 2. Research methodology flowchart



RESULTS AND DISCUSSION

Respondents and measurement model

Of 257 valid respondents, 38.5% were aged 35–44, 30.7% aged 25–34, 19.8% aged 45–54 and 10.9% over 54; gender was approximately balanced (51.4% female, 48.6% male). Trust in Platform had the lowest construct mean (M = 2.836) and Avoidance Motivation the highest (M = 3.800). On a single-choice item, 38.5% reported having activated Roblox's parental controls; the main reasons given by the others were not knowing how (18.7% of all respondents), believing the child was old enough (15.6%), finding the settings too complicated (10.5%) and doubting that the controls work (7.8%). Respondents who reported activation scored higher on Avoidance Behavior (3.50 vs 2.98), but agreement between the two measures was moderate (r = 0.27). All 28 indicators had outer loadings between 0.838 and 0.936, above the 0.70 threshold. Table 1 shows that all constructs satisfied convergent validity and reliability.

Table 1. Reliability and convergent validity

Construct	CA	CR	AVE
Perceived Threat	0.923	0.945	0.812
Trust in Platform	0.922	0.944	0.810
Safeguard Effectiveness	0.897	0.928	0.763
Safeguard Cost	0.947	0.962	0.863
Self-Efficacy	0.932	0.952	0.831
Avoidance Motivation	0.945	0.960	0.857
Avoidance Behavior	0.942	0.958	0.851

Note. CA = Cronbach's alpha; CR = composite reliability; AVE = average variance extracted.

Discriminant validity was confirmed by the Fornell-Larcker criterion and the HTMT ratio (Table 2). All 21 pairs fall below both the lenient 0.90 threshold and the stricter 0.85 threshold. The highest value, 0.695 (PT–AM), reflects the strong H3 structural path rather than a measurement problem. The two adjacent coping constructs, Safeguard Effectiveness and Self-Efficacy, have HTMT = 0.526, so they are empirically distinct although both depend on Trust in Platform. The full-collinearity VIFs, used to check common method bias, were all below 3.3

(1.481–2.324) (Kock, 2015). Inner-model VIFs ranged from 1.000 to 1.841, and item-level outer VIFs from 2.29 to 4.71, all below 5.

Table 2. Heterotrait-Monotrait (HTMT) matrix, all pairs below 0.85

	AB	AM	EF	PT	SC	SE
AM	0.658					
EF	0.316	0.048				
PT	0.514	0.695	0.244			
SC	0.062	0.350	0.581	0.541		
SE	0.223	0.076	0.526	0.096	0.327	
TP	0.168	0.343	0.495	0.463	0.436	0.480

Note. PT = Perceived Threat; TP = Trust in Platform; SE = Safeguard Effectiveness; SC = Safeguard Cost; EF = Self-Efficacy; AM = Avoidance Motivation; AB = Avoidance Behavior.

Structural model

The estimated model had an SRMR of 0.114 (saturated model: 0.037). Because SRMR has limited diagnostic value in variance-based PLS-SEM (Hair et al., 2022), we report this as a limitation rather than as grounds for rejecting the model. Avoidance Motivation (R^2 = 0.473) and Avoidance Behavior (R^2 = 0.387) show moderate explanatory power. Trust in Platform explains only about 20% of the variance in each coping construct (SE: R^2 = 0.195; EF: R^2 = 0.212), which suggests that other antecedents, such as digital literacy, parental age and peer influence, deserve attention. All Q^2 values exceeded zero. Tables 3 and 4 give the path and mediation results.

Table 3. Hypothesis testing results (bias-corrected bootstrapping, n = 5,000)

H	Path	β	t	p	f^2	Result
H1	TP → SE	0.441	9.202	< .001	0.241	Supported
H2	TP → EF	0.461	11.711	< .001	0.270	Supported
H3	PT → AM	0.626	14.973	< .001	0.548	Supported
H4	SE → AM	0.052	1.063	.288	0.004	Rejected
H5	SC → AM	+0.164	2.543	.011	0.028	Wrong direction
H6	EF → AM	0.244	3.843	< .001	0.067	Supported
H7	AM → AB	0.622	16.670	< .001	0.632	Supported

Note. PT = Perceived Threat; TP = Trust in Platform; SE = Safeguard Effectiveness; SC = Safeguard Cost; EF = Self-Efficacy; AM = Avoidance Motivation; AB = Avoidance Behavior.

Table 4. Specific indirect (mediation) effects (bias-corrected bootstrapping)

Mediation path	β	t	p	Result
TP → EF → AM	0.113	3.714	< .001	Significant
TP → SE → AM	0.023	1.070	.285	Not significant
TP → EF → AM → AB	0.070	3.534	< .001	Significant
TP → SE → AM → AB	0.014	1.063	.288	Not significant
PT → AM → AB	0.390	10.261	< .001	Significant
EF → AM → AB	0.152	3.656	< .001	Significant
SC → AM → AB	0.102	2.502	.012	Significant
SE → AM → AB	0.032	1.055	.292	Not significant

Note. PT = Perceived Threat; TP = Trust in Platform; SE = Safeguard Effectiveness; SC = Safeguard Cost; EF = Self-Efficacy; AM = Avoidance Motivation; AB = Avoidance Behavior.

Trust in Platform positively predicted Safeguard Effectiveness (H1: β = 0.441, p < .001) and Self-Efficacy (H2: β = 0.461, p < .001). Perceived Threat was the strongest predictor of Avoidance Motivation (H3: β = 0.626, p < .001, f^2 = 0.548). Safeguard Effectiveness did not significantly predict Avoidance Motivation (H4: β = 0.052, p = .288), so H4 was rejected. Safeguard Cost was positively rather than negatively associated with Avoidance Motivation (H5: β = +0.164, p = .011), the opposite of the hypothesized direction. Self-Efficacy predicted Avoidance Motivation (H6: β = 0.244, p < .001), and Avoidance Motivation strongly predicted Avoidance Behavior (H7: β = 0.622, p < .001, f^2 = 0.632). In the mediation analysis, Trust in Platform reached Avoidance Motivation through Self-Efficacy (β = 0.113, p < .001) but not through Safeguard Effectiveness (β = 0.023, p = .285). Five of the seven hypotheses were supported.

Two checks qualify these results. First, H6 depends on controlling for the other predictors. The zero-order correlation between Self-Efficacy and Avoidance Motivation is only r

= 0.04, while Self-Efficacy correlates negatively with Safeguard Cost (r = −0.55) and Perceived Threat (r = −0.23). The coefficient of 0.244 therefore emerges once PT and SC are held constant, which is a suppression pattern, and should be read as a conditional effect. Second, Trust in Platform has negative zero-order correlations with Avoidance Motivation (r = −0.32) and with Perceived Threat (r = −0.43). The positive paths from trust to the coping beliefs therefore do not mean that trust raises protective motivation overall.

DISCUSSION

Novelty. The main contribution of this study is to show, to our knowledge for the first time in the TTAT literature, how trust in the platform operates when the platform is both the source of the risk and the only provider of the safeguard. Previous TTAT and protection-motivation work has examined security behavior toward software the user installs, such as anti-malware or backup tools, or toward threats on devices the user controls (Blythe & Coventry, 2018; D. Q. Chen & Liang, 2019; McGill & Thompson, 2017; Sreenath et al., 2025; Thompson et al., 2017). We test platform trust inside this framework, in a child-safety setting where the user cannot choose another safeguard, and with parents acting on behalf of someone else.

Main findings. Five of seven hypotheses were supported. Trust in Platform positively affects both Safeguard Effectiveness (β = 0.441) and Self-Efficacy (β = 0.461), but only the self-efficacy route reaches motivation (β = 0.113 against 0.023 for effectiveness). Perceived Threat is the strongest driver of Avoidance Motivation (β = 0.626), and motivation strongly predicts behavior (β = 0.622). Two results run against the theory: Safeguard Effectiveness is unrelated to motivation (H4), and Safeguard Cost is positively related to it (H5). These findings matter because they show which of TTAT's assumptions hold, and which do not, when parents decide on behalf of their children.

Comparison with earlier studies. The dominant role of perceived threat matches the parent-based findings of Hwang et al. (2017), where perceived severity predicted both restrictive and active mediation, and Meng & Feng (2022), where perceived threat was the strongest predictor of privacy concern. Han et al. (2025) also found a positive threat effect, although response efficacy was larger there, and Sreenath et al. (2025) found that the TTAT model, which centers on perceived threat, explained more variance (about 64%) than PMT (about 60%). The role of self-efficacy agrees with the meta-analysis of Mou et al. (2022), in which self-efficacy and response efficacy had the largest effects on behavior, with Jansen & van Schaik (2018), who found self-efficacy among the strongest predictors of precautionary online behavior, with Glatz et al. (2018), who linked parental self-efficacy to internet-specific parenting practices, and with Alhelaly et al. (2024), who showed that self-efficacy conditions the move from threat appraisal to protection motivation. It differs from Sreenath et al. (2025), where self-efficacy was not supported in either model. The null effect of safeguard effectiveness is consistent with Thompson et al. (2017), who found that response efficacy did not affect intention in either the home computer or the mobile device group, but it contrasts with Hwang et al. (2017) and Jansen & van Schaik (2018), where response efficacy mattered, and with the strong response-efficacy effects summarized by Mou et al. (2022). For safeguard cost, Mou et al. (2022) found no support for response cost as a predictor of security motivation, and Sreenath et al. (2025) found it non-significant in both TTAT and PMT, whereas Blythe & Coventry (2018) and Han et al. (2025) reported the expected negative effect. We found a positive effect, which none of these studies reports. Finally, the strong motivation–behavior path (β = 0.622) is larger than what the intention-behavior gap literature would lead us to expect (Jenkins et al., 2021).

Why the results converge or differ. Several differences in design and context may explain the pattern. First, most comparison studies concern the user's own devices or accounts, where the person weighing a safeguard's effectiveness also bears the consequences. Protecting a child changes this: parental duty may override the evaluation of whether a tool works, which would explain why H4 was null here while response efficacy mattered in studies of self-protection (Jansen & van Schaik, 2018; Mou et al., 2022). We note that this is an interpretation of a null result, and we did not test it. Second, our parents have already passed an awareness screen, so variance in perceived effectiveness may be restricted. Third, the positive cost effect is small (f^2 = 0.028) and near the limit of detectability, and cross-sectional data cannot separate three

accounts: that complexity is read as a sign the safeguard is hard to bypass (cost-as-signal), that effort already spent inflates perceived value (effort justification), or that motivated parents simply meet more friction (reverse causality). Fourth, the self-reported behavior items may capture habitual self-description, which could inflate the H7 path relative to studies that observe behavior or that include a required-effort moderator (Jenkins et al., 2021). Fifth, trust may differ in meaning across populations: Khan et al. (2023) found that privacy cynicism lowered satisfaction but not trust among social media users, and Ayaburi & Treku (2020) found that trust in Facebook was restored through perceived integrity rather than reduced privacy concern. Both suggest that trust in a social platform may respond to different cues than the trust in lenders and banks that dominates the Indonesian studies we found (Pertiwi et al., 2025; Sebayang et al., 2024).

Contribution. Theoretically, the study extends TTAT with platform trust and shows that its influence runs through confidence rather than through beliefs about effectiveness, which is consistent with the affective-state source of self-efficacy (Bandura, 1977). It also sets out two boundary conditions for TTAT: when protection is exercised for someone else, effectiveness may stop acting as a gatekeeper (H4), and when the provider of the safeguard is also the source of the threat, an independent-safeguard assumption no longer holds. Empirically, it adds a PLS-SEM study of Indonesian parents on a user-generated-content platform, a setting that studies of Southeast Asian families have not covered, since they have focused on internet addiction and online risk exposure rather than on adoption of platform safety tools (Soh et al., 2018; Teimouri et al., 2018). The caveats reported in the Results section also add to the evidence: H6 is a conditional effect, and trust is negatively correlated with both threat perception and motivation at zero order.

Implications and future research. The findings point to a three-pillar policy ecosystem, although this is a synthesis of implications and was not tested directly. For regulators (Pillar 1), public communication under PP No. 17/2025 (Pemerintah Republik Indonesia, 2025) and PM Komdigi No. 9/2026 (Kementerian Komunikasi dan Digital Republik Indonesia, 2026) should rest on risk narratives that feel locally concrete, because Perceived Threat is the strongest motivational driver. Parents already rate the threat highly (M = 3.543), so the task is to make it concrete rather than to prove it exists. For schools (Pillar 2), voluntary parent workshops with follow-up checklists and open channels between schools and parents can narrow the gap between intention and action, provided they do not put children in the position of reporting on their parents. For platform developers (Pillar 3), a structured Indonesian-language channel through which parents can reach Roblox directly is a plausible way to raise Roblox's low trust score (M = 2.836, the lowest of any construct). Because trust seems to work through confidence, reducing the anxiety of using the controls may matter more than proving that they work. The same evidence points to feature and interface design. In the descriptive data, not knowing how and finding the settings too complicated outweighed doubts about effectiveness, which matches parents' reports elsewhere that parental controls are held back by usability gaps (Gnanasekaran & De Moor, 2025). Roblox added linked parent accounts, remote management and daily screen-time limits in late 2024 (Kaufman, 2024), yet respondents surveyed in April–May 2026 still reported these barriers, which suggests that adoption depends less on adding features than on making them easy to find and configure. Guided setup when a child account is created or linked, age-based presets, protective defaults (Information Commissioner's Office, 2020), in-app guidance in Indonesian and a dashboard that shows which controls are active follow from these results, although they were not tested here. Parents who mediate actively have also been found to favor both regulation and parental responsibility (Sharma & Lee, 2024), which suggests that regulatory messaging and parent-facing tools can be complementary. Future studies should use longitudinal or experimental designs to separate the three accounts of H5, measure actual activation of the controls, evaluate the parental-control interface directly through heuristic evaluation or usability testing, and test whether the trust paths replicate on other platforms.

Limitations

The most serious limitation is selection bias. Respondents needed some awareness of Roblox's parental controls to answer the effectiveness and self-efficacy items, which excludes parents with no awareness, arguably the least protected group, so the real adoption gap is

probably wider than observed. Second, all constructs came from one self-report instrument at one point in time, so common method variance may inflate the H7 path, even though the full-collinearity VIFs were acceptable (Kock, 2015; Podsakoff et al., 2024). Third, the cross-sectional design rules out causal claims for every path, including the two anomalous ones. Fourth, the sample is limited to Jabodetabek, where digital literacy, access and income may differ from other regions. Fifth, the findings are specific to Roblox and should not be assumed to transfer to TikTok, YouTube, Discord or Minecraft, since platform trust differs across platforms (Lee & Yuan, 2020). Sixth, activation status was a single self-reported item with only moderate agreement with the behavior scale, so the 38.5% figure is descriptive. Finally, the estimated-model SRMR (0.114) is above conventional cut-offs, and the power analysis shows that effects as small as H5 are close to the limit of detection.

CONCLUSION

This study tested a Trust-extended TTAT model of parental control adoption on Roblox with 257 Indonesian parents in Jabodetabek, and five of seven hypotheses held. Trust in Platform shapes both coping beliefs, but it reaches motivation through Self-Efficacy rather than Safeguard Effectiveness. Perceived Threat remains the strongest driver of motivation, and motivation strongly predicts self-reported behavior. The null H4 result is consistent with a stakes-override account and the positive H5 effect is best treated as a hypothesis for replication, not a conclusion; both need experimental tests. In practice, confidence-building communication from the platform, locally concrete risk messages from regulators and voluntary school-based workshops are the best-supported levers. Cross-platform replication would show how far the findings travel.

AUTHOR CONTRIBUTION

Wisnu Wardana: Conceptualization, Methodology, Data curation, Formal analysis, Writing – original draft. Ahmad Nurul Fajar: Supervision, Writing – review & editing.

DATA AVAILABILITY

Data supporting this study are not publicly available because of respondent confidentiality and can be requested from the corresponding author for legitimate academic purposes.

ACKNOWLEDGMENT

The authors thank the 257 parents in Jabodetabek who took part in the survey. The authors declare no conflict of interest. This research received no specific grant from any funding agency.

APPENDIX A: ITEM ADAPTATION

Table A1 shows how the 28-item instrument was contextualized to Roblox, with one representative adapted item and one original item per construct. TTAT items (PT, SE, SC, EF, AM, AB) were adapted from Liang & Xue (2010). Trust in Platform items (TP) were adapted from the trusting-beliefs scale of McKnight et al. (2002), because no earlier TTAT study had operationalized platform trust.

Table A1. Representative item adaptation from original sources to the Roblox context

Code	Roblox-adapted item	Original item
PT1–4	"My child is at high risk of experiencing negative outcomes (scams, bullying, adult content) while playing Roblox."	"Spyware poses a threat to me" (5 original items)
SE1–4	"Activating Roblox's parental control is effective in preventing my child from accessing inappropriate content."	"Anti-spyware software would be useful for detecting and removing spyware" (6 original items)
SC1–4	"Configuring Roblox's parental control requires a considerable amount of time."	"Installing anti-spyware software is too much trouble" (3 original items)

Code	Roblox-adapted item	Original item
EF1–4	"I am confident I can find and access Roblox's parental control settings menu."	"I could successfully install and use anti-spyware software if I had never used a package like it before" (10 original items)
AM1–4	"I strongly intend to use Roblox's available safety features to protect my child."	"I intend to use anti-spyware software to avoid spyware" (3 original items)
AB1–4	"I have activated parental control on my child's Roblox account."	"I run anti-spyware software regularly to remove spyware from my computer" (2 original items)
TP1–4	"I believe Roblox has strong technical capability to protect user safety." (Competence)	"LegalAdvice.com is competent and effective in providing legal advice" (Competence), adapted from McKnight et al. (2002); 10 original items

REFERENCES

Alhelaly, Y., Dhillon, G., & Oliveira, T. (2024). Mobile identity protection: The moderation role of self-efficacy. *Australasian Journal of Information Systems*, *28*, Article 18. https://doi.org/10.3127/ajis.v28.4397

Alpaullivarez, A. A., & Fajar, A. N. (2026). Exploring factors influencing user avoidance behavior toward QR-based payment scams. *Journal of Information Systems Engineering and Business Intelligence*, *12*(2), 458–473. https://doi.org/10.20473/jisebi.12.2.458-473

Altarturi, H. H. M., Saadoon, M., & Anuar, N. B. (2020). Cyber parental control: A bibliometric study. *Children and Youth Services Review*, *116*, Article 105134. https://doi.org/10.1016/j.childyouth.2020.105134

Anderson, C. L., & Agarwal, R. (2010). Practicing safe computing: A multimethod empirical examination of home computer user security behavioral intentions. *MIS Quarterly*, *34*(3), 613–643. https://doi.org/10.2307/25750694

Arsyad, I., & Fajar, A. N. (2026). Analyzing the determinants of intention to use property technology applications in Indonesia. *Journal of Mathematics Instruction, Social Research and Opinion*, *5*(3), 2399–2406. https://doi.org/10.58421/misro.v5i3.1332

Ayaburi, E. W., & Treku, D. N. (2020). Effect of penitence on social media trust and privacy concerns: The case of Facebook. *International Journal of Information Management*, *50*, 171–181. https://doi.org/10.1016/j.ijinfomgt.2019.05.014

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Blythe, J. M., & Coventry, L. (2018). Costly but effective: Comparing the factors that influence employee anti-malware behaviours. *Computers in Human Behavior*, *87*, 87–97. https://doi.org/10.1016/j.chb.2018.05.023

Carpenter, D., Young, A., Barrett, P., & McLeod, A. J. (2019). Refining technology threat avoidance theory. *Communications of the Association for Information Systems*, *44*, 462–487. https://doi.org/10.17705/1CAIS.04422

Chen, D. Q., & Liang, H. (2019). Wishful thinking and IT threat avoidance: An extension to the technology threat avoidance theory. *IEEE Transactions on Engineering Management*, *66*(4), 552–567. https://doi.org/10.1109/TEM.2018.2835461

Chen, H., Chen, Y., & Chen, J. (2024). Protecting teenagers' gaming privacy: The roles of parental mediation, platform protection, and risky encounters. *Behaviour & Information Technology*, *43*(14), 3578–3591. https://doi.org/10.1080/0144929X.2023.2285941

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage. https://collegepublishing.sagepub.com/products/research-design-5-255675

Du, Y., Grace, T. D., Jagannath, K., & Salen-Tekinbas, K. (2021). Connected play in virtual worlds: Communication and control mechanisms in virtual worlds for children and adolescents. *Multimodal Technologies and Interaction*, *5*(5), Article 27. https://doi.org/10.3390/mti5050027

ECPAT, INTERPOL, & UNICEF. (2022). *Disrupting harm in Indonesia*. ECPAT International. https://www.unicef.org/innocenti/media/4141/file/DH-Indonesia-Report-2022.pdf

Fajar, D. P., Mulyana, D., Dida, S., & Raksanagara, A. (2025). Parental digital illiteracy and communication paradox in children's video game practices. *Societies*, *15*(11), 317. https://doi.org/10.3390/soc15110317

Family Online Safety Institute & Ipsos. (2025). *Connected and protected: Insights from FOSI's 2025 online safety survey*. https://fosi.org/research/connected-and-protected-insights-from-fosis-2025-online-safety-survey/

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. *MIS Quarterly*, *27*(1), 51–90. https://doi.org/10.2307/30036519

Glatz, T., Crowe, E., & Buchanan, C. M. (2018). Internet-specific parental self-efficacy: Developmental differences and links to Internet-specific mediation. *Computers in Human Behavior*, *84*, 8–17. https://doi.org/10.1016/j.chb.2018.02.014

Gnanasekaran, V., & De Moor, K. (2025). Providing user perspectives to the next generation of parental controls. *International Journal of Child-Computer Interaction*, *44*, Article 100735. https://doi.org/10.1016/j.ijcci.2025.100735

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage. https://collegepublishing.sagepub.com/products/270548

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Han, M., Zhao, H., Ma, X., & Shi, R. (2025). Influencing factors of information security behavior among college students based on protection motivation theory: Evidence from China. *Frontiers in Public Health*, *13*, Article 1677024. https://doi.org/10.3389/fpubh.2025.1677024

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Huang, C. L., Yang, S. C., & Hsieh, L. S. (2019). The cyberbullying behavior of Taiwanese adolescents in an online gaming environment. *Children and Youth Services Review*, *106*, Article 104461. https://doi.org/10.1016/j.childyouth.2019.104461

Hwang, Y., Choi, I., Yum, J.-Y., & Jeong, S.-H. (2017). Parental mediation regarding children's smartphone use: Role of protection motivation and parenting style. *Cyberpsychology, Behavior, and Social Networking*, *20*(6), 362–368. https://doi.org/10.1089/cyber.2016.0555

Information Commissioner's Office. (2020). *Age appropriate design: A code of practice for online services*. https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/childrens-information/childrens-code-guidance-and-resources/age-appropriate-design-a-code-of-practice-for-online-services/

Jansen, J., & van Schaik, P. (2018). Testing a model of precautionary online behaviour: The case of online banking. *Computers in Human Behavior*, *87*, 371–383. https://doi.org/10.1016/j.chb.2018.05.010

Jenkins, J. L., Durcikova, A., & Nunamaker, J. F., Jr. (2021). Mitigating the security intention-behavior gap: The moderating role of required effort on the intention-behavior relationship. *Journal of the Association for Information Systems*, *22*(1), 246–272. https://doi.org/10.17705/1jais.00660

Kaufman, M. (2024, November 18). Major updates to our safety systems and parental controls. *Roblox*. https://about.roblox.com/newsroom/2024/11/major-updates-to-our-safety-systems-and-parental-controls

Kementerian Komunikasi dan Digital Republik Indonesia. (2026). *Peraturan Menteri Komunikasi dan Digital Nomor 9 Tahun 2026 tentang Tata Kelola Penyelenggaraan Sistem Elektronik dalam Pelindungan Anak*. https://jdih.komdigi.go.id/produk_hukum/view/id/1007/t/peraturan+menteri+komunikasi+dan+digital+nomor+9+tahun+2026

Khan, I., Loh, J. M. I., Hossain, A., & Talukder, M. J. H. (2023). Cynicism as strength: Privacy cynicism, satisfaction and trust among social media users. *Computers in Human Behavior*, *142*, Article 107638. https://doi.org/10.1016/j.chb.2022.107638

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of E-Collaboration*, *11*(4), 1–10. https://doi.org/10.4018/ijec.2015100101

Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal*, *28*(1), 227–261. https://doi.org/10.1111/isj.12131

Lee, Y.-H., & Yuan, C. W. (2020). The privacy calculus of "friending" across multiple social media platforms. *Social Media + Society*, *6*(2). https://doi.org/10.1177/2056305120928478

Liang, H., & Xue, Y. (2009). Avoidance of information technology threats: A theoretical perspective. *MIS Quarterly*, *33*(1), 71–90. https://doi.org/10.2307/20650279

Liang, H., & Xue, Y. (2010). Understanding security behaviors in personal computer usage: A threat avoidance perspective. *Journal of the Association for Information Systems*, *11*(7), 394–413. https://doi.org/10.17705/1jais.00232

Livingstone, S., & Helsper, E. J. (2008). Parental mediation of children's internet use. *Journal of Broadcasting & Electronic Media*, *52*(4), 581–599. https://doi.org/10.1080/08838150802437396

Livingstone, S., Ólafsson, K., Helsper, E. J., Lupiáñez-Villanueva, F., Veltri, G. A., & Folkvord, F. (2017). Maximizing opportunities and minimizing risks for children online: The role of digital skills in emerging strategies of parental mediation. *Journal of Communication*, *67*(1), 82–105. https://doi.org/10.1111/jcom.12277

McGill, T., & Thompson, N. (2017). Old risks, new challenges: Exploring differences in security between home computer and mobile device use. *Behaviour & Information Technology*, *36*(11), 1111–1124. https://doi.org/10.1080/0144929X.2017.1352028

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research*, *13*(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81

Meng, X., & Feng, B. (2022). Online taxi users' optimistic bias: China youths' digital travel and information privacy protection. *Frontiers in Psychology*, *13*, Article 1049925. https://doi.org/10.3389/fpsyg.2022.1049925

Mou, J., Cohen, J. F., Bhattacherjee, A., & Kim, J. (2022). A test of protection motivation theory in the information security literature: A meta-analytic structural equation modeling approach. *Journal of the Association for Information Systems*, *23*(1), 196–236. https://doi.org/10.17705/1jais.00723

Pemerintah Republik Indonesia. (2025). *Peraturan Pemerintah Nomor 17 Tahun 2025 tentang Tata Kelola Penyelenggaraan Sistem Elektronik dalam Pelindungan Anak*. https://jdih.komdigi.go.id/produk_hukum/view/id/965/t/peraturan+pemerintah+nomor+17+tahun+2025

Pertiwi, T. K., Joseph, C., Warmana, G. O., Khoirotunnisa, F., & Hariyana, N. (2025). Exploring platform trust, borrowing intention, and actual use of PayLater services in Indonesia and Malaysia. *Journal of Risk and Financial Management*, *18*(5), Article 255. https://doi.org/10.3390/jrfm18050255

Podsakoff, P. M., Podsakoff, N. P., Williams, L. J., Huang, C., & Yang, J. (2024). Common method bias: It's bad, it's complex, it's widespread, and it's not easy to fix. *Annual Review of Organizational Psychology and Organizational Behavior*, *11*, 17–61. https://doi.org/10.1146/annurev-orgpsych-110721-040030

Sciacca, B., Laffan, D. A., O'Higgins Norman, J., & Milosevic, T. (2022). Parental mediation in pandemic: Predictors and relationship with children's digital skills and time spent online in Ireland. *Computers in Human Behavior*, *127*, Article 107081. https://doi.org/10.1016/j.chb.2021.107081

Sebayang, T. E., Hakim, D. B., Bakhtiar, T., & Indrawan, D. (2024). What accelerates the choice of mobile banking for digital banks in Indonesia? *Journal of Risk and Financial Management*, *17*(1), Article 6. https://doi.org/10.3390/jrfm17010006

Sharma, S., & Lee, C. Y. (2024). Parental mediation and preferences for regulation regarding children's digital media use: Role of protection motivation and theory of planned behaviour. *Behaviour & Information Technology*, *43*(8), 1499–1517. https://doi.org/10.1080/0144929X.2023.2217275

Soh, P. C.-H., Chew, K. W., Koay, K. Y., & Ang, P. H. (2018). Parents vs peers' influence on teenagers' Internet addiction and risky online activities. *Telematics and Informatics*, *35*(1), 225–236. https://doi.org/10.1016/j.tele.2017.11.003

Sreenath, S. S. R., Hewitt, B., & Sreenath, S. (2025). Understanding security behaviour among healthcare professionals by comparing results from technology threat avoidance theory and protection motivation theory. *Behaviour & Information Technology*, *44*(2), 181–196. https://doi.org/10.1080/0144929X.2024.2314255

Statista. (2025). *Roblox: Hours spent by users under 13 in Q2 2025*. https://www.statista.com/statistics/1190610/hours-engaged-users-worldwide-roblox/

Teimouri, M., Benrazavi, S. R., Griffiths, M. D., & Hassan, M. S. (2018). A model of online protection to reduce children's online risk exposure: Empirical evidence from Asia. *Sexuality & Culture*, *22*(4), 1205–1229. https://doi.org/10.1007/s12119-018-9522-6

Tennakoon, H., Saridakis, G., & Mohammed, A.-M. (2018). Child online safety and parental intervention: A study of Sri Lankan internet users. *Information Technology & People*, *31*(3), 770–790. https://doi.org/10.1108/ITP-09-2016-0213

Thompson, N., McGill, T. J., & Wang, X. (2017). "Security begins at home": Determinants of home computer and mobile device security behavior. *Computers & Security*, *70*, 376–391. https://doi.org/10.1016/j.cose.2017.07.003

UNICEF Indonesia. (2023). *Online knowledge and practice of parents and children: Indonesia baseline study 2023*. https://www.unicef.org/indonesia/media/23586/file/online-knowledge-practice-parents-and-children-Indonesia-baseline-study-2023.pdf

Wang, G., Zhao, J., Van Kleek, M., & Shadbolt, N. (2021). Protection or punishment? Relating the design space of parental control apps and perceptions about them to support parenting for online safety. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2). https://doi.org/10.1145/3476084

Wisniewski, P. J., Ghosh, A. K., Xu, H., Rosson, M. B., & Carroll, J. M. (2017). Parental control vs. Teen self-regulation: Is there a middle ground for mobile online safety? *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 51–69. https://doi.org/10.1145/2998181.2998352