

Accessibility and Deceptive Design in Parental Control Features - Barriers to Effective Caregiver Supervision

Raquel S. Santos

raquel.silvas@ufrpe.br

Departamento de Estatística e Informática, UFRPE
Recife, Pernambuco, Brazil

George Valença

george.valenca@ufrpe.br

Departamento de Computação, UFRPE
Recife, Pernambuco, Brazil

ABSTRACT

Digital platforms increasingly provide parental control features to support caregivers in supervising children's online experiences. However, the mere availability of such features does not guarantee that caregivers can effectively locate, understand, and configure them. Our study investigates whether parental control interfaces contain accessibility barriers and deceptive design patterns that may undermine effective supervision. We conducted a systematic inspection protocol of 15 parental-control configuration flows comprising 150 screens across TikTok, Instagram, Discord, and YouTube. The analysis combined accessibility criteria derived from ABNT NBR 17060, ABNT NBR 17225, and WCAG 2.2 with deceptive pattern taxonomies proposed by Gray et al. (2018, 2024) and the Organisation for Economic Co-operation and Development (OECD). To support cross-platform comparison, we operationalized three analytical indicators that capture cognitive load, configuration friction, and the co-occurrence of accessibility barriers and deceptive design. The results reveal that accessibility limitations and deceptive patterns frequently appear together in parental-control interfaces, creating cumulative interaction barriers that hinder caregiver supervision. Based on these findings, we developed a set of medium-fidelity prototypes grounded in accessibility standards and the Child Rights by Design framework. Such redesign illustrates how identified barriers can be mitigated through more transparent, accessible, and caregiver-centered interfaces. These findings suggest that parental controls alone are insufficient for effective protection and highlight the need for transparency, accessibility, and rights-oriented design practices aligned with children's rights and regulations.

KEYWORDS

Parental Control, Deceptive Patterns, Accessibility, Child Safety

1 INTRODUCTION

Digital platforms have become a primary environment for children's social, educational, and recreational life. In Brazil, the TIC Kids Online 2025 survey [8] indicates that 92% of children and adolescents aged 9 to 17 are Internet users, corresponding to approximately 24.5 million individuals. The study also highlights the central role of digital platforms in young people's daily lives. Among Internet users in this age group, WhatsApp is accessed several times a day (53%), followed by YouTube (48%), Instagram (48%), and TikTok (46%). In parallel, other platforms, such as Discord, continue to

attract adolescents through interest-based communities and social interaction features [24]. This broad and early exposure to digital environments has increased pressure on families, regulators, and platform companies to ensure that children's online experiences are safe, age-appropriate, and effectively supervised by adults.

In response to growing societal and regulatory pressure, major platforms have introduced parental control features that allow adults to configure restrictions, monitor activity, and manage aspects of a young user's digital experience. In Brazil, this movement has been reinforced by Lei 15.211/2025 (ECA Digital), which requires platforms to provide intuitive and transparent parental supervision tools [4]. Similar regulatory initiatives have emerged in the European Union [11, 12], the UK [23, 30], and the United States [31], reflecting a growing international consensus that platform design plays a central role in protecting children online.

Yet the mere presence of parental control features does not guarantee their effective use. Interface design plays a critical role in determining whether caregivers can successfully locate, understand, and configure available controls, or whether the setup process itself becomes a barrier to supervision. Therefore, the existence of a feature does not necessarily translate into meaningful protection in practice. Human-Computer Interaction research shows that caregivers frequently encounter difficulties when interacting with parental control interfaces [13, 25], particularly those with lower levels of digital literacy [35]. As a result, regulatory frameworks that mandate the provision of parental control mechanisms without addressing their usability may fail to ensure that these tools effectively support children's online safety.

This gap is further reinforced by research documenting the widespread adoption of deceptive design patterns across popular mobile applications [7, 10, 14]. These patterns, which denote interface strategies that exploit human behavior to steer users toward actions that primarily benefit the platform rather than the user, are increasingly recognized by regulators and policymakers as a distinct source of consumer harm [22]. Despite this growing concern, limited attention has been given to whether such patterns also shape the configuration processes caregivers must navigate to activate parental supervision features. If parental control interfaces incorporate deceptive design, they may undermine caregivers' ability to exercise the protections that platforms formally make available.

In light of this context, we performed an analytical inspection of parental control features across four major platforms: TikTok, Instagram, Discord, and YouTube. Configuration flows were evaluated against two complementary frameworks: accessibility standards derived from ABNT NBR 17060, ABNT NBR 17225, and WCAG 2.2, and a deceptive-pattern typology based on Gray et al. [14, 15] and the OECD [22]. These standards provide established criteria for

perceivability, operability, understandability, and robustness, while also addressing factors such as navigation clarity, predictable behavior, and readable content that influence usability and cognitive effort. This framework enables the assessment of various caregiver profiles, including disabled users, older adults, and those with limited digital literacy. We further investigate whether accessibility barriers and deceptive design are not independent phenomena in parental control interfaces, particularly whether conditions that hinder navigation also facilitate deceptive elements. Based on these findings, we developed medium-fidelity prototypes illustrating how accessibility standards and Child Rights by Design principles can mitigate these barriers through improved interface design.

2 CONCEPTUAL BACKGROUND

In this section, we present the three conceptual pillars that underpin the study: restrictive parental mediation in the digital environment (Section 2.1), manipulative design as a barrier to user autonomy (Section 2.2), and accessibility standards as a minimum condition for the use of tools (Section 2.3). The articulation between these factors is essential for understanding the results described in the subsequent sections.

2.1 Parental Control on Digital Platforms

Parental control refers to the set of technical mechanisms and configurations that enable caregivers to manage children's digital experiences on platforms and devices. It has increasingly been understood as a broader process of digital education and guidance. Drawing on evidence from the Brazilian context, Bueno et al. [6] argue that effective parental mediation encompasses not only monitoring and restriction practices, but also active dialogue, shared use of digital media, and the development of children's critical understanding of online environments. Their analysis highlights that digital education is a continuous effort requiring caregivers to support children in interpreting online experiences, assessing risks, and developing autonomous and responsible digital behaviors.

Within digital platforms, parental control features represent one of the primary mechanisms through which this supervisory role is operationalized. These features are intended to provide caregivers with practical means to manage children's online experiences by configuring restrictions, monitoring activities, and controlling access to platform functionalities. As a field of practice, it encompasses account linking between responsible adults and supervised minors, content filtering and age-based restrictions, management of social interactions and contacts, screen time and schedule controls, and activity monitoring and reporting. These mechanisms translate the broader concept of restrictive parental mediation into configurable features within platform interfaces [17].

Research on the practical implementation of parental controls has consistently documented a gap between the formal availability of these tools and their actual use, from a socio-technical perspective. As argued by Wang et al. [33], parental control applications vary substantially in transparency, granularity of restrictions, and support for parent-child communication, factors that directly influence caregivers' willingness to adopt these technologies. This means that their effectiveness depends not only on the scope of the

controls offered but also on how these controls are integrated into the user experience and made accessible to caregivers.

2.2 Deceptive Designs and Barriers

According to Brignull [5], interface strategies that exploit human behavior to induce actions favoring the platform over the user represent deceptive design patterns. These strategies have been systematically documented across popular mobile applications. Gray et al. [14] organized these patterns into five macro-classes: *Nagging* (persistent interruptions), *Obstruction* (structural barriers), *Sneaking* (concealing information), *Interface Interference* (manipulating visual hierarchy), and *Forced Action* (requiring unrelated tasks). This taxonomy was expanded by Di Geronimo et al. [10] into sixteen sub-classes via a qualitative expert inspection protocol, which has since been replicated internationally, such as by Hidaka et al. [16].

In the Brazilian context, Coelho et al. [7] identified recurrent forms of navigational friction and interface-induced decision bias in mobile applications used by Brazilian users, demonstrating how interaction flows can subtly influence user behavior beyond explicit persuasion mechanisms. Likewise, Neris et al. [19] discussed how usability barriers may operate as structural constraints that disproportionately affect vulnerable user groups, particularly when important configuration or privacy-related actions are embedded within complex interaction paths.

Among the documented patterns, *Obstruction* and *Interface Interference* are particularly consequential for the usability of parental control features. Obstruction manifests in parental control contexts as configuration options buried within multi-level submenus, absent from primary navigation menus, or reachable only through non-obvious paths, imposing navigational costs that deter caregivers before they can activate supervision. Interface Interference manifests as commercial engagement elements, promotional notifications, and platform-specific prompts that compete for caregiver attention during the configuration process, redirecting it away from the supervision task.

The OECD [22] formalized these dynamics in a regulatory framework defining dark commercial patterns as practices that create unnecessary friction in users' decision-making processes—a characterization that directly implicates platform design as a mechanism of accountability rather than merely a matter of usability. Gray et al. [15] then developed an ontology of deceptive patterns synthesizing definitions, relationships, and measurement approaches across the existing literature. The complete typology of criteria adopted in this study follows this approach and will be presented later.

2.3 Accessibility and Ethical Design

The normative antithesis of deceptive and obstructive design is a framework of accessible and ethical design principles that center user comprehension, autonomy, and informed decision-making as design requirements rather than secondary considerations. Digital accessibility, as defined by the W3C Web Accessibility Initiative [32], refers to the set of design properties that enable an interface to be perceived, understood, navigated, and used by different user profiles, regardless of their individual characteristics, abilities, or level of familiarity with the digital environment. The international

normative framework is grounded in the Web Content Accessibility Guidelines (WCAG 2.2), which organizes requirements under four principles: perceivable, operable, understandable and robust. These principles define measurable success criteria referenced by legislation and standards in numerous jurisdictions.

Two Brazilian standards implement these principles locally: ABNT NBR 17060 [2] establishes accessibility requirements for mobile applications, while ABNT NBR 17225 [3] extends them to web content, aligned with WCAG 2.2. Both standards demand features critical for parental control interfaces, such as semantic labeling, plain language, adequate contrast, contextual help, visible feedback, predictable navigation and appropriate information density. Despite these regulations, Ceweb.br/NIC.br [20] documented that, as of 2024, only 2.9% of Brazilian websites comply with accessibility standards [18], establishing the scale of the gap between normative requirements and actual platform practice.

Research on the cognitive dimensions of interface complexity provides a complementary basis for evaluating parental control flows. Cognitive Load Theory, developed by Sweller [26] and expanded in subsequent foundational works [27, 28], distinguishes between intrinsic load, determined by the task's inherent complexity, and extraneous load, which is generated by unnecessary design elements. Sweller's principle of element interactivity establishes that designs which increase the number of interactive elements beyond what the task requires impose unnecessary cognitive costs on users. Zhang and Son [34] applied these principles to mobile interfaces, documenting a significant relationship between element density per screen and cognitive demand.

These technical frameworks align with a broader normative agenda for child protection in digital environments. The 5Rights Foundation's Child Rights by Design (CRbD) framework translates children's rights into eleven design principles, notably *Agency* [1]. This principle evaluates whether products support the free, informed choices of caregivers and children, or deploy features that undermine caregiver autonomy [1], directly challenging interfaces that conceal consequences, delegate initiative to minors, or embed commercial elements in configuration flows. Together, ABNT standards, cognitive load principles, and the 5Rights framework establish the evaluative axis used to interpret this study's findings.

3 RESEARCH METHOD

This study is guided by the following research question: *to what extent do the parental control features offered by digital platforms popular among Brazilian youth provide the necessary conditions of accessibility and transparency for restrictive parental mediation to be effectively exercised?* To address this question, the research was structured into three sequential phases: **scope definition**, **data collection** and **computational analysis**. The procedures are briefly described in Figure 1.

3.1 Definition of Scope

Platform selection followed three sequential criteria. "First, we identified, based on the *frequency of daily use* among Brazilian youth aged 9–17, WhatsApp (53%), YouTube (48%), Instagram (48%), and TikTok (46%) as the most frequently used platforms [8]. Second, applying *functional relevance* (requiring native parental controls

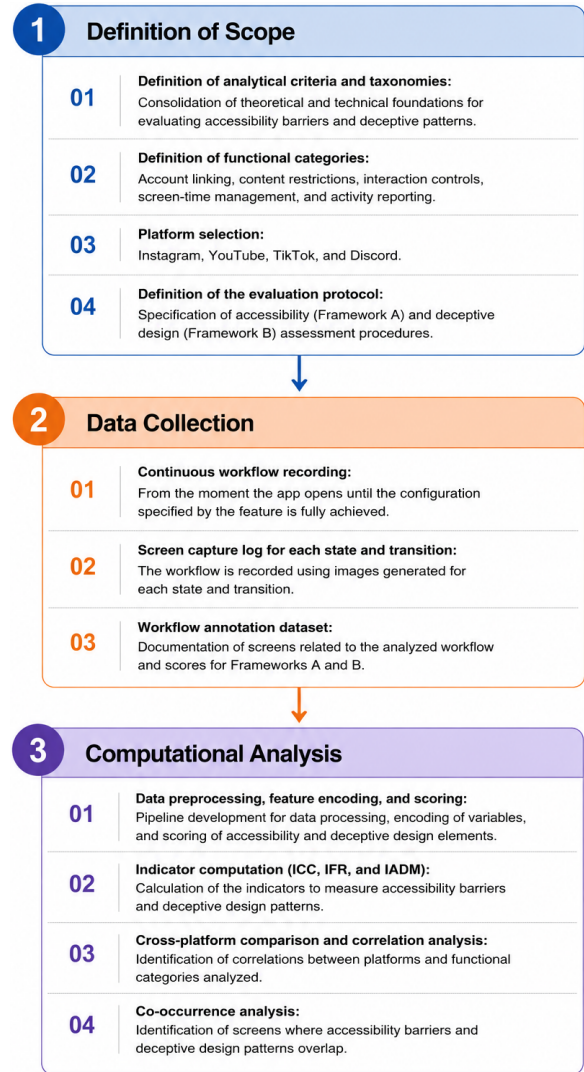


Figure 1: Summary of the procedures followed for the analysis of the corpus.

across at least some of our five functional categories) excluded WhatsApp due to its messaging-centric architecture lacking a distinct supervision module. Conversely, Discord (8%) was included because it offers a dedicated family-supervision interface and architectural variance from Meta's ecosystem, whereas Facebook (19%) was excluded to avoid infrastructural redundancy with Instagram.

The third criterion, *platform diversity*, ensured that the final selection represented three distinct architectures of parental-control implementation: group-based supervision, dedicated native modules and distributed settings across a community-based interface. This structural variety enables systematic comparison of how design choices within each implementation model affect the usability of parental supervision features.

3.2 Data Collection

Following an expert-inspection protocol adapted from Di Geronimo et al. [10], data were collected through structured manual inspection of the selected platforms. The protocol was adapted to focus on parental control configuration flows, incorporate accessibility criteria, and mitigate classification bias and pattern-recognition fatigue in a single-researcher inspection. Accordingly, the process was divided into two temporally separated phases — data capture and evaluation — to reduce DP blindness [10], i.e., the tendency to overlook deceptive elements during interface interaction.

Operationally, a standardized task script was developed for each platform to assess entry points, action sequences, and success criteria for feature configuration. All interactions were screen-recorded, with screenshots documenting interface states and transitions. Each screen was assessed against two criteria sets: (A) accessibility, addressing semantic accessibility, language clarity, contrast, contextual help, action labelling, feedback, navigation consistency, and information density; and (B) deceptive patterns, addressing hidden or misleading information, interface interference, bundling or forced consent, obstruction, and nagging. Results were recorded as binary classifications with qualitative observations in a structured dataset.

3.3 Computational Analysis

The binary inspection records were processed computationally through an analytical pipeline that converts per-screen observations into three scalar indicators per configuration flow, enabling systematic cross-platform and functional-category comparison.

Let f denote a configuration flow, N_f the total number of screens traversed in f , and i an individual screen index. For each screen i , two aggregate scores are computed: an *accessibility score* $a_i \in [0, 1]$ — the proportion of criteria A1–A8 satisfied — and a *deceptive pattern score* $b_i \in [0, 1]$ — the proportion of criteria B1–B6 present. The three indicators are formally defined as follows.

3.3.1 Cognitive Load Index (CLI). Grounded in Cognitive Load Theory [28] — which establishes that human working memory has limited capacity and that design elements beyond task requirements impose unnecessary cognitive burden —, the CLI is proposed because it captures the compounding effect of depth and density on cognitive load as the product of flow depth (number of screens) and mean interactive element density (average elements per screen).

$$CLI_f = P_f \times \bar{Q}_f \quad (1)$$

where P_f is the number of screens traversed in flow f and $\bar{Q}_f$ is the arithmetic mean of interactive elements per screen. Higher CLI values indicate longer, denser configuration paths that impose greater demands on working memory [28], increasing the likelihood of caregiver error, disorientation, or task abandonment.

3.3.2 Friction Index (FI). The FI aggregates three structural penalties that impose configurability barriers independently of the caregiver's digital literacy:

$$FI_f = 2|B5_f| + \max(0, P_f - \tilde{P}) + \delta_f \quad (2)$$

where $|B5_f|$ is the count of screens in flow f presenting an obstruction pattern (criterion B5), weighted by 2; P_f is the depth of

the flow; $\tilde{P}$ is the median depth across the corpus; and $\delta_f \in \{0, 1\}$ equals 1 if the flow requires an initial action from the supervised minor — a design choice that subordinates caregiver control to the minor's cooperation [17].

These principles are grounded in foundational work on task-centered interface design and empirical methods for measuring user experience, both of which establish that configurability and task completion are measurable dimensions of interface quality [29]. The obstruction component is weighted by 2, as it represents a categorical usability failure, not merely impaired performance. Users faced with non-obvious navigation paths abandon tasks at a rate approximately twice that of users navigating through transparent, albeit longer, paths of equivalent depth [21]. This weighting also reflects the fact that obstruction directly violates Guideline 2.4 (Navigability) of WCAG 2.2 in two ways: by compromising both predictability (users cannot anticipate where the options are) and detectability (the options are not detectable through standard navigation).

3.3.3 Accessibility and Manipulative Design Index (AMD). The Accessibility and Manipulative Design Index (AMD) was developed to investigate whether inaccessibility and deceptive design co-occur systematically on the same screens, suggesting cumulative interaction effects that compound user harm beyond independent violations. It is expressed as the Spearman rank-order correlation between the per-screen accessibility score and deceptive pattern score across all screens comprising a flow or platform:

$$AMD_f = (\rho_f, r_f) \quad (3)$$

where ρ_f is the *deceptive pattern density* — the proportion of screens in flow f presenting at least one of the criteria (B1–B6) — and r_f is the *Spearman correlation* between the per-screen accessibility score $a_i \in [0, 1]$ and the per-screen deceptive pattern score $b_i \in [0, 1]$ across all screens i in f . A negative r_f indicates that more manipulative screens tend to exhibit lower accessibility. Spearman's rank-order coefficient was selected because both a_i and b_i are ordinal proportions derived from binary variables [9].

The pipeline is illustrated in Figure 2: binary per-screen observations are aggregated by flow, yielding the screen-level scores a_i and b_i , from which CLI, FI, and AMD are derived according to Equations 1–3.

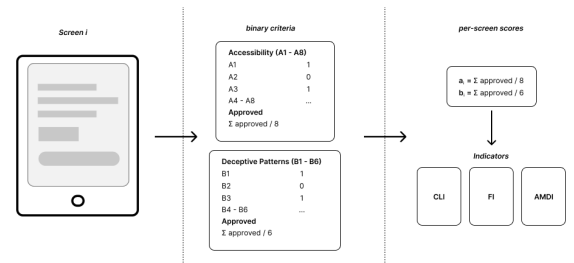


Figure 2: Summary of the processes individually applied to the screens of the analyzed workflows for a comprehensive analysis of the workflow based on mapped metrics.

4 RESULTS

This section presents the results of the analysis of parental control features across the four platforms based on the three proposed indicators. The corpus comprises 15 workflows and 150 screens distributed across five functional categories. We interpret these findings in terms of what they reveal about the conditions platforms provide (or fail to provide) for caregivers to exercise effective restrictive mediation. Table 1 presents a comparative overview of the corpus across platforms.

The global Spearman correlation ($r_s = 0.328$) between accessibility and deceptive design suggests an association between accessibility and deceptive design scores. Screens with more deceptive elements tend to present greater accessibility barriers, suggesting that both factors may reinforce each other and jointly hinder caregivers' ability to understand and configure parental controls.

4.1 TikTok

TikTok comprises 42 screens across four of the five functional categories investigated, excluding activity reporting, which was unavailable. It exhibits the highest density of deceptive patterns in the study. Intra-platform analysis further shows that these patterns are unevenly distributed, with deceptive elements and accessibility barriers frequently co-occurring on the same screens. This is particularly exposed on entry screens, where caregivers may encounter both types of barriers at the earliest stages of configuration.

The account-pairing workflow, which serves as the entry point for all other parental supervision features, illustrates this pattern. The interface combines the absence of textual labels, contextual help, and visual feedback with distracting elements that promote native platform features and provide limited guidance for locating supervisory settings (Figure 3). Additionally, the workflow makes recurring use of untranslated English terms and platform-specific jargon — as in "Push", "Live" and "Add" — increasing the cognitive effort required to understand and configure parental controls.

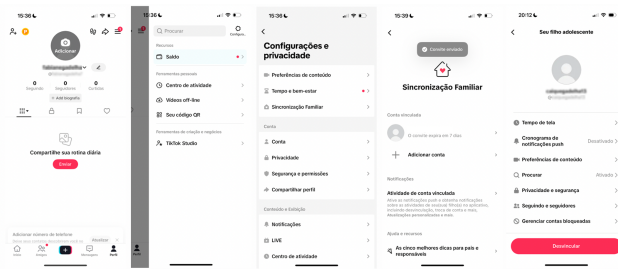


Figure 3: Initial workflow for parental control settings.

4.2 Instagram

Instagram formally offers the broadest set of parental supervision features among the analyzed platforms, covering all five functional categories across 60 screens. However, this breadth is accompanied by the highest concentration of configuration barriers and greater structural complexity. The account-pairing workflow alone spans five levels of depth and 17 screens, with some of the highest densities of interactive elements in the corpus (Figure 4). These findings

indicate that a broader set of supervisory tools does not necessarily translate into greater accessibility or ease of use for caregivers.

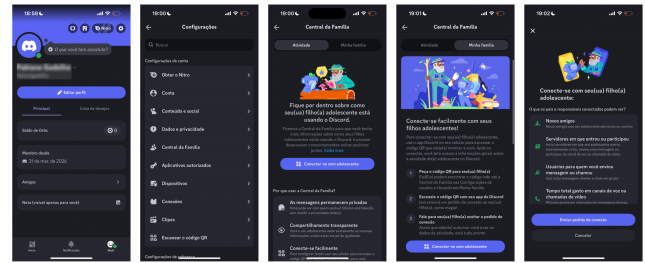


Figure 4: Inter-account pairing flow.

For content and interaction restrictions, configuration options are often nested within multiple layers of menus organized around specific limitations. In several cases, the irreversible consequences of certain restrictions are not disclosed before the action is executed (Figure 5). Furthermore, the absence of immediate feedback throughout the process leaves caregivers without clear confirmation that their actions have been successfully applied. Together, these findings suggest a mismatch between the breadth of supervisory functionality offered by the platform and the clarity with which these features can be understood and configured.

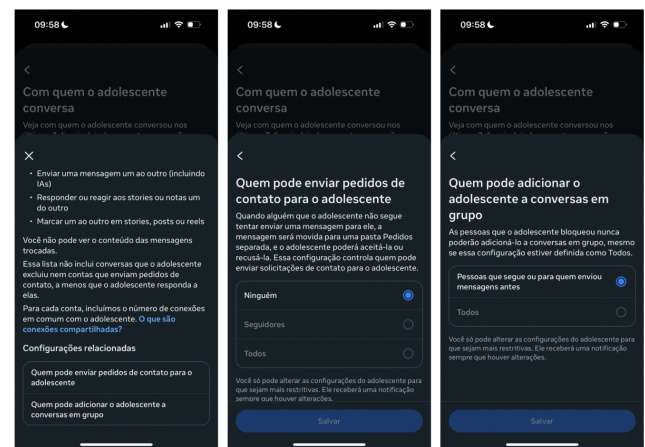


Figure 5: Limitations on settings communicated at the point of action.

4.3 Discord

Discord was analyzed across four functional categories, with screen-time management features being the only functionality not identified, resulting in a corpus of 30 screens. Although the platform presents the lowest average cognitive load among the analyzed platforms, the nature of its barriers is qualitatively distinct and carries specific implications for caregivers' ability to provide informed supervision. In particular, the strongest intra-platform correlation was associated with the simple language criterion (A2), which was

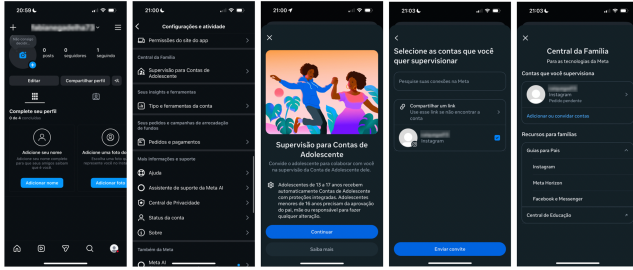
Table 1: Overview of the corpus by platform.

Platform	Flows	Total Screens	Avg. CLI	Avg. FI	Deceptive Patterns Density	Avg. Accessibility Score	Avg. Deceptive Patterns	r_s (Spearman)
TikTok	4	42	82.0	4.2	0.541	0.860	0.135	-0.401
Instagram	5	60	129.4	10.4	0.506	0.873	0.129	-0.490
Discord	4	30	72.5	6.5	0.395	0.763	0.120	-0.572
YouTube	2	18	87.5	0.0	0.275	0.853	0.073	0.325
Overall	15	150	96.0	6.3	0.455	0.837	0.120	-0.328

satisfied in only 36.7% of screens (the lowest compliance rate observed for any criterion across all platforms in the study).

This situation reflects a recurring reliance on English terminology and platform-specific vocabulary (such as “Get Nitro”, “Server Boost” and “Orb Amount”) without sufficient contextualization for users unfamiliar with the platform. The issue is reinforced by the repeated use of such terminology throughout the configuration process, increasing the effort required to understand supervisory settings and controls. Furthermore, the activation of parental supervision must be initiated by the child on their own device.

This design choice introduces a dependency on the minor’s participation, potentially limiting the caregiver’s ability to independently configure restrictive mediation features. In parallel, Obstruction (B5) was identified in 40% of the analyzed screens, primarily through the absence of clear pathways to parental resources within home screens, settings menus, and submenus. Examples of these issues are illustrated in Figure 6.

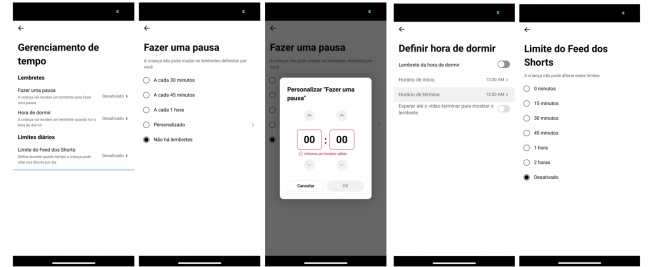
**Figure 6: Application of specific terms to the platform without context.**

4.4 Youtube

YouTube was analyzed across two workflows (account pairing and screen-time management) totaling 18 screens. Features for content restriction, interaction restriction, and activity reporting were not available within the platform’s parental supervision interface. Hence, YouTube offers a substantially narrower range of parental mediation features than the other platforms analyzed. This limited scope is an important finding in its own right, as the availability of supervision mechanisms directly shapes the extent to which caregivers can monitor and manage a child’s digital experience.

Across both workflows, YouTube exhibited the lowest density of deceptive patterns in the corpus, reflecting a comparatively direct and streamlined configuration process. Its workflows illustrate how parental supervision features can be implemented through a limited number of clearly signaled steps, with greater transparency and navigational clarity.

However, YouTube was the only platform that showed a positive Spearman correlation between accessibility and scores for misleading patterns. This suggests a specific risk: misleading elements may appear on screens that are otherwise well-structured and accessible, which can make it more difficult to detect misleading information. Therefore, it is important to analyze the indicators in their entirety to investigate the implications of this aspect. For example, contextual help (A4) was present on only 50% of screens, matching the cross-platform average and potentially limiting support during decision-making. Additionally, the screen-time management workflow presents reminders as time-control mechanisms without clearly explaining their operation or scope. Thus, caregivers may activate features without fully understanding their limitations (Figure 7), demonstrating that navigational clarity alone does not ensure transparent communication of functionality.

**Figure 7: Workflow for creating reminders when usage time is reached.**

5 DESIGN IMPLICATIONS AND PROTOTYPING

Across the four platforms analyzed, the availability of parental control features did not necessarily translate into effective support for parental mediation. The inspection revealed persistent gaps between the formal provision of supervision tools and the conditions required for caregivers to locate, understand, and configure these functionalities effectively. Although deceptive design strategies vary across platforms, they commonly limit informed parental control by making supervision tools more difficult to access, obscuring the consequences of configuration choices, misrepresenting the scope of available controls, or diverting caregivers’ attention away from parental supervision features. As a result, these strategies impose additional cognitive and operational burdens on caregivers, undermining the effectiveness of parental mediation.

The inspection protocol and comparative indicators contribute methodologically by enabling systematic identification and comparison of accessibility barriers and deceptive design patterns across

platforms and functional categories. Our findings also provide empirical evidence for assessing the extent to which parental control interfaces support informed parental mediation, offering a basis for the development of design guidelines, evaluation frameworks, and regulatory requirements aligned with accessibility principles and children's rights frameworks.

The CRbD framework, together with the accessibility standards established by ABNT NBR 17225 and WCAG 2.2, provides a normative foundation for addressing these shortcomings. Accordingly, this section presents medium-fidelity prototype comparisons based on the inspected interfaces, placing original screens alongside redesigned versions that address the identified issues. Each comparison connects the observed problem to its corresponding accessibility or child-rights criterion and illustrates the proposed design intervention. Thus, the recommendations are grounded in established accessibility requirements and children's rights obligations rather than design preferences.

The first practice concerns **locus of control**: caregivers should retain full autonomy over the activation, configuration, and deactivation of supervision features, without depending on the child's actions. This requires clear communication of who controls each setting, explicit confirmation mechanisms, and interfaces focused on the caregiver's task rather than platform engagement. The TikTok Search configuration screen illustrates this failure: two radio buttons ("Enabled" and "Disabled") were presented without describing the controlled feature, identifying who could modify it, or providing a save or confirmation mechanism. As shown in Figure 10, adding feature-scope information, action-oriented labels, caregiver exclusivity, and explicit save control addresses criteria A1, A2, A4, A5, and IFR.

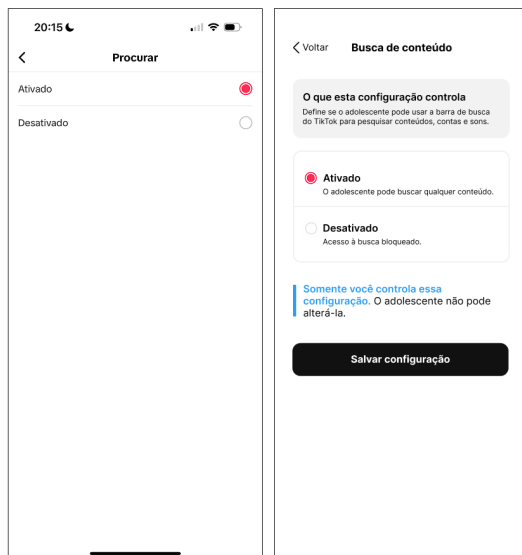


Figure 8: Prototype comparison between the TikTok Search configuration screen and a suggested adaptation featuring adjacent scope descriptions, action-labeled options, a caregiver-exclusivity notice, and an explicit save button.

The second practice concerns **transparency of scope**: the actual scope and limitations of each feature should be communicated before the caregiver confirms a setting, rather than discovered after activation. This includes disclosing irreversible settings, non-relaxable restrictions, and consequences for the supervised account, while labelling interactive elements according to their resulting actions rather than using generic terms. The Instagram location-sharing screen illustrates this failure: the toggle appeared active while the effective condition—that the adolescent was not sharing location with anyone—was presented separately below, creating ambiguity about what the control governed. As shown in Figure 9, stating the current state and feature limitations before the control and integrating them into the decision flow addresses criteria B1, B2, and A4.

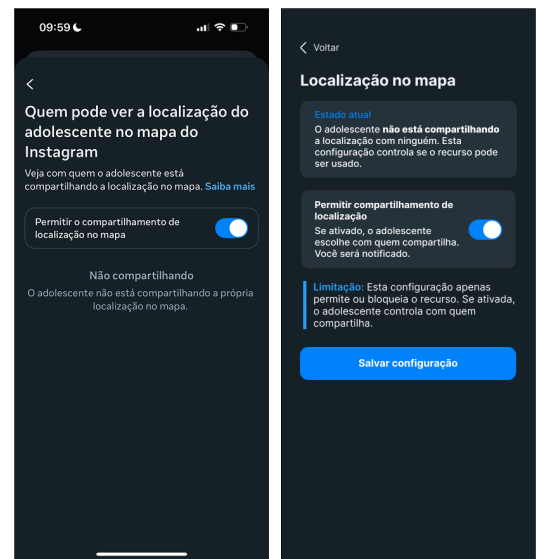


Figure 9: Prototype comparison between Instagram's location-sharing screen and suggestions for improvements to provide a realistic view of the feature's controls and limitations.

The third practice concerns the **configuration path**: supervision features should be accessible through a clearly signposted entry point, with navigation depth proportional to feature complexity and consistent cues such as return paths, progress indicators, and clear exit points. In the inspected corpus, Youtube's supervision listed categories without showing the configuration state, requiring caregivers to open each submenu to verify progress. The absence of progress indicators or distinctions between completed and pending items further increased this burden. Figure 8 shows how displaying configuration status alongside each menu item and adding a progress indicator addresses criteria A7, A8, and B5, transforming the overview into an actionable supervision dashboard.

The fourth practice concerns **contextual support**: each configuration option should provide a plain-language description of its function, scope, and limitations at the decision point, without

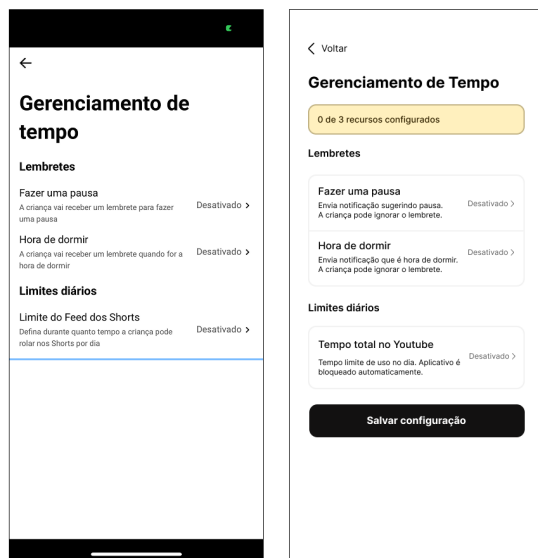


Figure 10: Prototype comparison between the YouTube Kids Time Management screen and a suggested adaptation featuring plain-language descriptions, explicit feature limitations, a progress indicator, and configuration state confirmation.

requiring external navigation. Visible and explicit feedback should also confirm that each action has been successfully registered. This aligns with the contextual help requirements of ABNT NBR 17225 and transparency principles in children’s rights frameworks. In the inspected corpus, however, options were often presented with generic labels and no adjacent explanation, while links such as “learn more” redirected caregivers away from the decision point. Figure 11 illustrates this issue in Discord’s teen account settings and shows how integrating descriptions alongside each toggle, without external redirection, addresses criteria A2, A4, and A6.

6 CONCLUSION

The aim of this study was to investigate whether parental control features on popular digital platforms provide the accessibility and transparency necessary for parents and guardians to exercise effective restrictive mediation. The findings reveal a consistent gap between the formal availability of these features and the conditions required for their effective use. Although all analyzed platforms offer some form of parental supervision, accessibility barriers, opaque configuration processes, and deceptive design elements frequently hinder caregivers’ ability to locate, understand, and configure available controls.

As a **contribution**, we present a reproducible analysis protocol and three operationalized indicators that translate accessibility and deceptive design criteria into a structured assessment tool. Accordingly, we propose design recommendations to address critical issues across platforms and improve user experience. The results reframe parental control flaws as rights-based obligations, shifting the focus from the mere availability of supervision tools to their effective

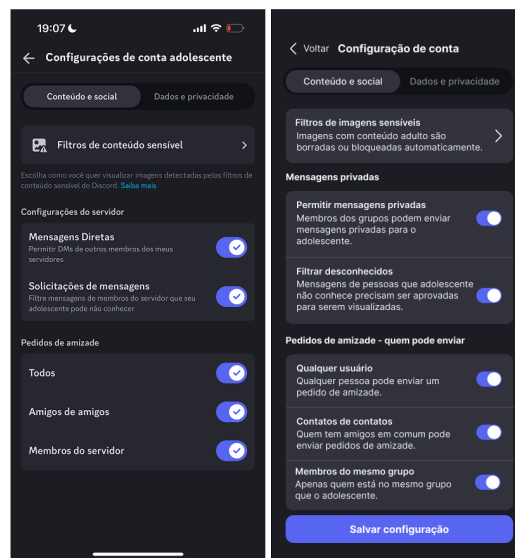


Figure 11: Prototype comparison between Discord’s Content and Social screen and suggested improvements providing contextual support for content, message, and friend request filters.

usability, a reframing that guides the implementation of the ECA Digital and analogous regulatory frameworks.

The main **limitation** is that the inspection was conducted by a single researcher, introducing potential classification inconsistency. This risk is mitigated by binary operationalized criteria and mandatory qualitative descriptions, but cannot be eliminated. The limited corpus also restricts cross-platform comparability for some functional dimensions. Moreover, the indicators assess interface structure rather than caregiver experience, requiring empirical validation with caregivers to establish their relationship with supervision outcomes.

In **future studies**, we should focus on calibration of the indicators through studies involving caregivers with different levels of digital literacy, enabling the establishment of empirical thresholds in parental control workflows. Additionally, collaborating with UI and accessibility experts will help verify normative criteria and clarify interface consequences. Finally, the observed convergence between accessibility standards and deceptive design patterns should be further explored in collaboration with regulatory bodies, supporting the translation of these findings into verifiable design requirements for platform accountability.

7 ETHICAL CONSIDERATIONS

Not applicable. This study did not involve human participants or the collection of personal data from individuals. Given these circumstances, no participant consent procedure was conducted and no participant-identifying information was included in the research dataset. The evidence analyzed consisted of interface states, screenshots, screen recordings, and structured observations generated during the inspection of the selected platforms and were restricted to information to document interface structure, interaction sequences and accessibility characteristics.

8 ACKNOWLEDGMENTS

The authors acknowledge the use of generative artificial intelligence tools exclusively for translation support. All content was reviewed and validated by the authors.

REFERENCES

- [1] 5Rights Foundation and Digital Futures Commission. 2023. Child Rights by Design. 5Rights Foundation. <https://childrightsbydesign.5rightsfoundation.com>
- [2] Associação Brasileira de Normas Técnicas. 2022. ABNT NBR 17060: Acessibilidade em aplicativos para dispositivos móveis — Requisitos. ABNT.
- [3] Associação Brasileira de Normas Técnicas. 2025. ABNT NBR 17225: Acessibilidade em conteúdo e aplicações web — Requisitos. ABNT.
- [4] Brasil. 2025. Lei n. 15.211, de 2 de setembro de 2025: Dispõe sobre a proteção de crianças e adolescentes em plataformas digitais (ECA Digital). Diário Oficial da União. <https://www.planalto.gov.br>
- [5] Harry Brignull. 2023. *Deceptive Patterns: Exposing the Tricks Tech Companies Use to Control You*. Testimonium Ltd, London. <https://www.deceptive.design>
- [6] Bernardo Bueno, Laura Magalhães Moreira, Juliana Potter, Aline Restano, and Daniel Tornaim Spritzer. 2025. Digital Education is a Constant Effort: Reflections on Parental Mediation in Children's Use of Social Media. In *Survey on Internet Use by Children in Brazil: ICT Kids Online Brazil 2024*. Cetic.br, Brazil. <https://bibliotecadigital.acervo.nic.br/entities/publication/cf7da32a-988a-4482-b2e8-4d8b8f587e01>
- [7] Amanda Coelho, Gabriel Lima, Lucas Melo, Levi Ribeiro, and Ticianne Darin. 2025. Shades of Manipulation: A Large-Scale Analysis of Deceptive Design in Brazilian Mobile Applications. In *Anais do XXIV Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais* (Belo Horizonte/MG). SBC, Porto Alegre, RS, Brasil, 1157–1182. doi:10.5753/ihc.2025.10905
- [8] Comitê Gestor da Internet no Brasil. 2025. *TIC Kids Online Brasil 2025: Pesquisa sobre o uso da Internet por crianças e adolescentes no Brasil*. Technical Report. NIC.br, São Paulo. <https://cetic.br/pt/pesquisa/kids-online/>
- [9] Gregory W. Corder and Dale I. Foreman. 2014. *Nonparametric Statistics: A Step-by-Step Approach* (2nd ed.). John Wiley & Sons, Hoboken, NJ. doi:10.1002/9781118142486
- [10] Linda Di Geronimo, Larissa Braz, Enrico Fregnan, Fabio Palomba, and Alberto Bacchelli. 2020. UI Dark Patterns and Where to Find Them: A Study on Mobile Applications and User Perception. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3313831.3376600
- [11] European Commission. 2025. Commission Publishes Guidelines on the Protection of Minors under the Digital Services Act. Accessed: 2026-05-27. <https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-protection-minors>
- [12] European Union. 2022. Digital Services Act - Article 28 Protection of Minors. Accessed: 2026-05-27. https://www.eu-digital-services-act.com/Digital_Services_Act_Article_28.html
- [13] Arup Kumar Ghosh, Karla Badillo-Urquiola, Shion Guha, Joseph J. LaViola Jr, and Pamela J. Wisniewski. 2018. Safety vs. Surveillance: What Children Have to Say about Mobile Apps for Parental Control. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3173574.3173698
- [14] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. 2018. The Dark (Patterns) Side of UX Design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3173574.3174108
- [15] Colin M. Gray, Cristiana Teixeira Santos, Nataliia Bielova, and Thomas Mildner. 2024. An Ontology of Dark Patterns Knowledge: Foundations, Definitions, and a Pathway for Shared Knowledge-Building. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 289, 22 pages. doi:10.1145/3613904.3642436
- [16] Shun Hidaka, Sota Kobuki, Mizuki Watanabe, and Katie Seaborn. 2023. Linguistic Dead-Ends and Alphabet Soup: Finding Dark Patterns in Japanese Apps. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 3, 13 pages. doi:10.1145/3544548.3580942
- [17] Sonia Livingstone and Ellen Helsper. 2008. Parental Mediation of Children's Internet Use. *Journal of Broadcasting Electronic Media* 52 (12 2008), 581–599. doi:10.1080/08838150802437396
- [18] Movimento Web para Todos. 2024. *Pesquisa de Acessibilidade Digital: sites brasileiros 2024*. Technical Report. Movimento Web para Todos, São Paulo. <https://mwpt.com.br>
- [19] Vânia P. A. Neris, Jean Clemisson Santos Rosa, Cristiano Maciel, Vinicius Carvalho Pereira, Vinicius F. Galvão, and Isabela L. Arruda. 2024. GrandIHC-BR 2025-2035 - GC4: Sociocultural Aspects in Human-Computer Interaction. In *Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems (IHC '24)*. Association for Computing Machinery, New York, NY, USA, Article 20, 14 pages. doi:10.1145/3702038.3702057
- [20] NIC.br / Ceweb.br. 2023. Todos na Web: Conferência de Acessibilidade Digital. NIC.br. <https://todosnaWeb.nic.br>
- [21] J. Nielsen. 2000. *Designing Web Usability: The Practice of Simplicity* (1st ed.). New Riders Publishing.
- [22] Oecd. 2022. *Dark commercial patterns*. OECD Digital Economy Papers 336. OECD Publishing. doi:10.1787/44f5e846-en
- [23] Ofcom. 2025. Protecting Children from Harms Online. Accessed: 2026-05-27. <https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/statement-protecting-children-from-harms-online>
- [24] Pew Research Center. 2024. Teens and Video Games Today. (May 2024). Accessed: 2026-08-20. <https://www.pewresearch.org/internet/2024/05/09/teens-and-video-games-today/>
- [25] Irwin Reyes, Primal Wijesekera, Joel Reardon, Amit Elazari, Abbas Razaghpahan, Narseo Vallina-Rodriguez, and Serge Egelman. 2018. "Won't Somebody Think of the Children?" Examining COPPA Compliance at Scale. *Proceedings on Privacy Enhancing Technologies* 2018 (04 2018), 63–83. doi:10.1515/popets-2018-0021
- [26] John Sweller. 1988. Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science* 12, 2 (1988), 257–285. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1207/s15516709cog1202_4 doi:10.1207/s15516709cog1202_4
- [27] John Sweller. 2010. Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. *Educational Psychology Review* 22 (06 2010), 123–138. doi:10.1007/s10648-010-9128-5
- [28] John Sweller, Jeroen J. G. Van Merriënboer, and Fred Paas. 2019. Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review* 31 (06 2019), 261–292. doi:10.1007/s10648-019-09465-5
- [29] Thomas Tullis and William Albert. 2008. *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics: Second Edition*. Morgan Kaufmann.
- [30] UK Government. 2023. Online Safety Act 2023. Accessed: 2026-05-27. <https://www.gov.uk/government/publications/online-safety-act-explainer/online-safety-act-explainer>
- [31] United States Senate. 2024. Kids Online Safety Act. Accessed: 2026-05-27. <https://nypost.com/2024/07/30/business/senate-passes-landmark-legislation-for-childrens-online-safety/>
- [32] W3C Web Accessibility Initiative. 2024. Introduction to Web Accessibility. World Wide Web Consortium (W3C). <https://www.w3.org/WAI/fundamentals/accessibility-intro/>
- [33] Ge Wang, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2021. Protection or Punishment? Relating the Design Space of Parental Control Apps and Perceptions about Them to Support Parenting for Online Safety. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 343 (Oct. 2021), 26 pages. doi:10.1145/3476084
- [34] Ruiqi Zhang and Oo-Sung Son. 2025. A Measurement Model for Visual Complexity in HCI: Focusing on Visual Elements in Mobile GUI Design. *Electronics* 14 (02 2025), 942. doi:10.3390/electronics14050942
- [35] Pengfei Zhao, Natalie Bazarova, and Natercia Valle. 2023. Digital parenting divides: the role of parental capital and digital parenting readiness in parental digital mediation. *Journal of Computer-Mediated Communication* 28 (08 2023). doi:10.1093/jcmc/zmad032