

Beyond the Feed: How Generative AI is Reshaping Children's Social Media Ecologies in the Global South

Nermeen Singer

nermin.singer@chi.asu.edu.eg

Ain Shams University

Fathey El Sharkawy

Ain Shams University

Amira Ismail

Canadian International College

Article

Keywords: generative AI, social media, children, Global South, Egypt, datafication, platform power, cultural bias

Posted Date: September 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10317656/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Social media platforms are no longer simple communication channels for the world's children and teenagers: they are increasingly infrastructures within which generative artificial intelligence (AI) operates as a constant, often invisible layer, curating feeds, generating content, and, through AI companions embedded in mainstream apps, even simulating relationships. Scholarship on children and social media has developed largely independently of scholarship on children and generative AI, even though the two now converge inside the same applications young people use every day. This article proposes a conceptual bridge between these two literatures through the concept of double mediation: the idea that children's platform experiences are shaped twice over, first by the datafied logics of the platform itself, and second by the generative AI systems layered on top of it, which carry their own cultural and linguistic assumptions. We develop and apply this framework to the case of Egypt and the wider Arab region, drawing on demographic data, the growing body of research on cultural and dialectal bias in Arabic-language AI systems, and the region's distinct religious and family-mediation practices. We then situate the framework within the current global wave of social-media age-restriction policy, using Australia's Social Media Minimum Age Act and Egypt's own emerging legislative debate as contrasting cases, to argue that regulatory frameworks focused solely on age verification address only half of the mediation children now experience. We close with a research agenda for scholars working at the intersection of childhood studies, platform studies, and AI research, particularly from Global South perspectives.

1. Introduction

When scholars first began mapping how social media platforms shape the everyday lives of children and teenagers, the object of study was comparatively legible: a feed, a friend list, a set of algorithmic recommendations pulling from a known, if opaque, set of signals. Concerns centred on identity performance, peer belonging, cyberbullying, commercial exploitation, and the slow erosion of a clear line between childhood and the attention economy. These concerns remain entirely valid, and are indeed the ones motivating the current wave of regulatory activity this Collection engages directly, from Australia's world-first under-16 restriction to the parallel debates now underway in Britain, France, and elsewhere.

But the platforms themselves have changed underneath this scholarship. Generative AI is no longer a separate application that children might choose to open alongside their social media use; it has become infrastructural to the platforms themselves. Snapchat's "My AI" sits inside the same interface as a teenager's group chats. Meta has built AI-driven companion characters that can message users proactively across Instagram, Facebook and WhatsApp, at times without making clear to the user that the account is not a human being. Recommendation algorithms that once merely selected among existing human-made content increasingly rank, and now generate, synthetic content directly. For a child or teenager opening a social media app in 2026, the practical distinction between "using social media" and "using generative AI" has largely dissolved.

This article argues that the scholarly literatures on children-and-social-media and on children-and-generative-AI have developed largely on separate tracks, and that the moment has come to draw them together, not as two adjacent risks to be listed side by side, but as a single, compounding process of mediation. We propose the concept of double mediation to describe this process: children's experience of social platforms is shaped first by the datafied logic of the platform (what is collected, what is optimised for, what is shown), and shaped a second time by the generative AI layer now built into that platform, which carries its own, often Western-centric, linguistic, cultural and representational assumptions.

We develop this framework with particular attention to children in the Global South, and specifically in Egypt and the wider Arab region, for two reasons. First, the empirical and policy literature on children, social media and AI remains disproportionately generated in, and about, North American and European contexts, even as the platforms in question operate globally and even as some of the largest national populations of young social media users live outside those regions. Second, the double mediation we describe is not evenly distributed: children whose language, dialect, religious practice and family structures differ from the training data and design assumptions embedded in dominant AI systems experience a qualitatively different, and under-examined, version of platform life.

The article proceeds in four further parts. We first review the separate literatures on children's social media use and children's engagement with generative AI, identifying where they converge and where that convergence has so far gone unremarked. We then set out the double mediation framework in full. We apply the framework to the Egyptian and Arab regional case, drawing on demographic data, recent computational-linguistics literature on Arabic-language AI bias, and the distinct religious and family-mediation context that shapes how Egyptian children and their caregivers negotiate online life. Finally, we turn to policy, reading the current global wave of age-restriction legislation, including Egypt's own emerging proposals, against the framework, and arguing that a regulatory focus limited to age verification addresses only one half of what children now navigate online.

2. Literature Review

2.1 From Audiences to Platforms: Children in the Age of Social Media

The dominant theoretical vocabulary for understanding how platforms organise social life comes from what van Dijck, Poell and de Waal (2018) call the platform society: an ecosystem in which platforms are fuelled by data, organised by algorithms, and governed through business models and user agreements that most users never meaningfully read. Three mechanisms, in this account, give platforms their power: datafication, the translation of virtually any activity into data that can be processed and monetised; commodification, the transformation of that data, and of users' attention, into tradable value; and selection, the algorithmic sorting that determines what any individual user sees. Applied to children specifically, this framework has produced a substantial body of work on what Lupton and Williamson

(2017) term the datafied child: a young person increasingly rendered, from birth, as what they memorably describe as an algorithmic assemblage, tracked and profiled by both commercial platforms and, increasingly, by the institutions around them.

Livingstone and Third (2017) reframed this datafication in explicitly rights-based terms, arguing for an emerging agenda in which children's digital lives are understood not only through a protectionist lens of risk and harm but through the fuller set of rights, to participation, provision and protection, that children hold both online and offline. This rights framing has since been extended by Mascheroni and Siibak (2021), whose account of datafied childhoods traces how children across a range of national contexts imagine, narrate and sometimes resist the data practices that surround them, and by Stoilova, Livingstone and Nandagiri (2020), who show that children's capacity to understand and manage their own data and privacy depends heavily on digital literacy support from schools and peers, support that is unevenly available.

What unites this literature, for present purposes, is that it treats the platform, and the data-driven selection mechanism at its core, as the primary object of concern. Generative AI barely features, for the simple reason that most of this scholarship precedes the point at which generative AI became a default, embedded feature of the platforms it studies. The empirical and conceptual tools this literature has built, particularly its attention to datafication and to children's rights as a framing device, remain indispensable. But they were not built to account for a platform that does not merely select among existing content but actively generates new content, and new simulated relationships, in response to a child's input.

A closely related strand of this scholarship has examined identity performance and peer culture directly, showing how children and teenagers use the affordances of a given platform, filters, disappearing messages, curated grids, to experiment with self-presentation in front of an imagined peer audience, often navigating multiple, context-specific versions of themselves across different platforms and different segments of their social network. This literature has consistently found that platform design choices, far from being neutral backdrops, actively shape which versions of identity are easiest to perform and which are structurally discouraged, an insight that anticipates, without yet incorporating, the argument this article develops about generative AI's own role in narrowing the identities available to a child at the point of content creation itself.

2.2 Generative AI Enters the Picture: From Tool to Companion

A parallel and, until recently, largely separate literature has tracked children's and teenagers' direct engagement with generative AI tools. Early work in this area, including this article's own contribution to a wider programme of research on media psychology and children's engagement with digital and screen-based media (Singer, 2023a, 2023b, 2024, 2025), tended to frame AI as a tool: something used for homework help, creative writing, image generation, or entertainment, broadly analogous to earlier waves of research on children's use of search engines, educational software, or, in still earlier media

environments, animated cartoons as a vehicle for information-processing development and incidental vocabulary acquisition (Singer, 2019, 2022).

That framing has been overtaken by events. Reporting from Common Sense Media and Stanford Medicine's Brainstorm Lab found that a large majority of surveyed teenagers, roughly seven in ten in the United States, had used an AI chatbot for companionship at least once, with a substantial minority reporting that their conversations with an AI companion felt as satisfying as, or more satisfying than, conversations with real people (Vasan et al., cited in Andoh, 2025). Researchers posing as teenagers were able to elicit harmful dialogue, on subjects including self-harm, sexual content and violence, from several widely used companion applications with very little prompting, and documented cases in which chatbots failed to intervene, or actively validated, expressions of suicidal ideation. These findings, and a small number of associated wrongful-death lawsuits, prompted rapid, if uneven, regulatory and platform responses through late 2025 and into 2026, including new parental controls and teen-specific behavioural limits from OpenAI, Meta and Character.AI, California's first-in-the-nation companion-chatbot safety statute (SB 243, effective January 2026), an inquiry by the US Federal Trade Commission into several major AI companies, and transparency obligations for conversational AI under the EU AI Act due to apply from August 2026.

Crucially for this article's argument, several of the applications named in this literature, Snapchat's My AI foremost among them, are not standalone companion apps but features built directly into mainstream social media platforms already central to children's peer lives. A recent survey study of adolescents' use of My AI found that engagement with the chatbot, and adolescents' emotional reactions to it, varied by age, gender and socioeconomic status, and noted that as children's capacity to distinguish an AI system from a human interlocutor develops with age, so too does the emotional register of their engagement with it. This is not, in other words, a separate technology children opt into; it is a feature that arrives already installed inside the platforms this Collection is concerned with.

It would be a mistake, however, to read this literature as uniformly alarmist. The same generative AI features that raise the safety concerns documented above are also, for many children and teenagers, genuinely participatory and creative tools: engines for fan fiction, remix and meme culture, avatar and image generation, and collaborative storytelling that extend, rather than replace, the creativity and play the Collection's own call identifies as a distinct area of scholarly interest. The concern this article raises is not that generative AI features are inherently harmful, but that both their risks and their creative affordances are currently evaluated almost entirely against a template built from Anglophone, Western-market use, leaving open the question of how well either the risk profile or the creative affordance actually holds for children operating in a different linguistic and cultural register.

What this second literature has so far under-examined, by contrast with the platform studies tradition reviewed above, is the structural, datafied, culturally-situated character of the systems it studies. The dominant frame is psychological and safety-oriented: what does this technology do to an individual child's wellbeing, attachment style, or risk exposure? Comparatively little of this literature asks the

platform-studies question of whose data, whose language, and whose cultural assumptions these companion systems are built from, and what that means for children outside the markets in which these systems were designed and tested. It is precisely this gap, between a platform-studies literature attentive to datafication but not to generative AI, and an AI-safety literature attentive to generative AI but not to platform-level datafication and cultural power, that the following section addresses.

3. Conceptual Framework: Double Mediation

We propose double mediation as a framework for understanding how children now experience AI-embedded social platforms, one that draws the two literatures reviewed above into a single analytic model rather than treating them as parallel concerns.

The first mediation is the one platform-studies scholarship has already theorised in detail: the datafication, commodification and selection processes described by van Dijck, Poell and de Waal (2018). A child's activity on a platform, what they view, for how long, what they type, who they interact with, is continuously converted into data, that data is put to commercial use, and an algorithmic selection process determines what that child, and every other user, subsequently sees. This first mediation is now well understood to be neither neutral nor evenly distributed: it privileges content and behaviour that sustains engagement, and it encodes the commercial priorities of the platform operator rather than the interests of the child user, a point at the heart of the rights-based critique developed by Livingstone and Third (2017) and by Lupton and Williamson (2017).

The second mediation is newer, and is the one this article seeks to bring into the same frame. Generative AI systems, whether embedded as a chat companion, a caption generator, a content-moderation classifier, or a recommendation engine that now creates rather than merely selects content, are themselves trained on data that is neither neutral nor evenly representative of the world's languages and cultures. A growing body of computational linguistics research has documented this unevenness in detail. Naous, Ryan, Ritter and Xu (2024) show, using a large dataset of naturally occurring Arabic-language prompts, that both multilingual and Arabic-monolingual language models default to entities, foods, clothing, and even religious framings associated with Western culture over comparable Arab equivalents, a bias that becomes more pronounced, counterintuitively, the more linguistically fluent and natural the Arabic prompt. Related work on dialectal variation within Arabic itself finds that this bias is not uniform even across Arabic-speaking populations: lower-resource dialects, including Iraqi Arabic, show measurably stronger negative stereotyping in model outputs than higher-resource varieties such as Modern Standard Arabic or Gulf Arabic (AraRegBias, 2026). Efforts to build Arabic-centric alternatives, including the Jais and Jais-Chat family of models and ArabianGPT, are explicit acknowledgements from within the AI research community that the dominant, English-centric training pipeline systematically underserves the roughly 300 million-plus speakers of Arabic worldwide, and that Arabic's dialectal and morphological complexity is easily flattened by approaches optimised for English.

Double mediation names the compounding effect of these two processes operating on the same child, inside the same application, often within the same few seconds of use. A platform's selection mechanism has already narrowed what a child sees toward whatever sustains engagement; a generative AI layer sitting on top of that selection, whether generating a caption suggestion, an AI companion's reply, or a translated or moderated version of a friend's post, then narrows it a second time toward whatever the AI system's own training distribution treats as the cultural default. The two processes are not simply additive. Because generative AI outputs are themselves a new, highly engaging form of content, they feed back into the first mediation's datafication and selection loop: a child's positive engagement with a Western-coded AI companion response becomes a data point that shapes what the platform's selection algorithm subsequently promotes, further embedding the cultural defaults of the second mediation into the first.

It is worth situating double mediation briefly against neighbouring theoretical vocabulary, to clarify what the concept adds. It is related to, but distinct from, the idea of the algorithmic imaginary, the ways users come to imagine and adapt to an algorithm's presumed logic even without direct knowledge of its workings: double mediation describes an objective, two-layer technical process, not primarily a user's subjective mental model of it, though the two are certainly connected in practice. It is also related to, but broader than, the datafied-child literature's account of children as algorithmic assemblages (Lupton & Williamson, 2017): where that account emphasises the accumulation of data traces over a child's life course, double mediation emphasises the compounding of two distinct, layered technical systems, platform selection and generative AI, operating on a single interaction in something closer to real time. The two framings are complementary rather than competing, and a full account of datafied childhood in the generative AI era will likely need both the longitudinal, trace-accumulation view and the layered, real-time view this article foregrounds.

This framework makes two claims that the remainder of the article develops. The first is diagnostic: existing children's-rights and datafication scholarship, however sophisticated its account of the first mediation, cannot fully explain a child's platform experience without also accounting for the second. The second is distributive: because the second mediation's cultural defaults are not universal but are skewed toward the languages and cultural contexts best represented in AI training data, double mediation is experienced unevenly across the world's children, in ways that map closely, though not perfectly, onto existing global inequalities between the Global North and Global South. We now turn to a case that makes this distributive claim concrete.

4. The Global South Case: Egypt and the Arab Region

4.1 Digital Childhoods in Egypt: Scale and Shape

Egypt offers a case study in which the scale of children's platform engagement, the specificity of the linguistic and cultural context, and the emergence of domestic policy debate all converge. Internet penetration in Egypt reached 81.9 percent of the population by early 2025, with approximately

96.3 million individual users, and roughly 50.7 million social media user identities, a figure that grew by nearly 12 percent in a single year (DataReportal, 2025). Instagram alone reported some 20.1 million users in Egypt in early 2025, equivalent to a substantial share of the platform's eligible audience aged 13 and above, while TikTok's user base grew sharply over the same period. These figures describe a national population in which social media is not a marginal youth subculture but a near-default mode of communication, entertainment and information for anyone with a smartphone.

Within that population, national research bodies have begun to document the specific exposure of children. A 2024 report by Egypt's National Center for Social and Criminological Research, a government-affiliated think tank, found that approximately half of children under eighteen in Egypt use social media platforms in ways that expose them to harmful content, cyberbullying or abuse, a finding that was cited directly by Egyptian lawmakers in early 2026 as they began drafting legislation to regulate children's social media use following a public call from President Abdel Fattah el-Sisi to address what parliamentary statements described as “digital chaos” facing the country's children. Egypt's emerging policy debate, in other words, closely tracks the international wave of concern that motivates this Collection, but it does so from within a linguistic and cultural context substantially different from the Anglophone contexts in which most of the underlying platform and AI research has been conducted.

Two further features of the Egyptian digital landscape are relevant to how double mediation plays out in practice. First, access is overwhelmingly mobile-first: the great majority of Egyptian internet users, including children and teenagers, encounter social platforms primarily through smartphones rather than through the desktop environments in which much of the earlier children's-internet-safety literature was developed, a pattern that shapes both the pace of scrolling-based engagement with AI-generated content and the practical difficulty of the kind of sustained parental co-viewing that desktop-era mediation research often assumed. Second, Egypt's own connectivity infrastructure, while improving rapidly, with median mobile download speeds rising from roughly 22 to 24 megabits per second between 2024 and 2025 alone, still lags substantially behind Gulf peers, meaning that Egyptian children's platform experience is shaped not only by the cultural and linguistic dynamics discussed below but by a distinct, and still evolving, material infrastructure of access.

This digital expansion has taken place alongside a broader national push to align Egyptian education with the United Nations' sustainable development agenda, itself framed around measurable targets for digital literacy and future-oriented skills (Singer & El-Sayed Mahmoud, 2020). That comparative policy work, tracking the gap between Egypt's stated 2020 education priorities and its 2030 sustainable development targets, provides useful context for the present discussion: an Egyptian child's encounter with AI-embedded social platforms is unfolding inside an education system that is itself still building the digital and media-literacy infrastructure that earlier scholarship (Stoilova, Livingstone, & Nandagiri, 2020) identifies as a precondition for children's meaningful understanding of, and agency over, the data and AI systems that mediate their platform use.

4.2 Language, Dialect, and Cultural (Mis)recognition

It is at this point that the double mediation framework does its analytic work. An Egyptian child's first mediation, the platform-level datafication and selection process, operates on infrastructure built by the same handful of American technology companies that dominate the platform society globally (van Dijck et al., 2018). Its second mediation, however, is shaped by the specific, documented weaknesses of generative AI systems when handling Arabic, and Egyptian Arabic in particular. Naous et al. (2024) demonstrate that this is not a marginal technical limitation but a systematic bias: language models asked to complete naturally occurring Arabic prompts about food, clothing, names, or beverages default to Western-associated entities at rates that increase as the Arabic prompt becomes more fluent and idiomatic, precisely the register in which an Egyptian teenager would naturally write. Related research on Arabic dialectal bias finds that Egyptian Arabic, while comparatively better resourced than some other regional dialects, remains far less well represented in AI training data than Modern Standard Arabic or English, and that stereotype-amplifying behaviour in generative outputs varies significantly by dialect (AraRegBias, 2026). The practical consequence is that an Egyptian child using an AI-generated caption suggestion, an AI companion's conversational reply, or an AI content-moderation decision on a colloquial Egyptian Arabic post, is more likely than a child working in English or Modern Standard Arabic to encounter a system that misreads, flattens, or defaults away from their own linguistic and cultural register.

This is not an abstract concern. The same body of research that documents rising use of AI companions among Anglophone teenagers, discussed in Section 2.2, describes systems explicitly designed to build a sustained conversational relationship with the user, remembering prior exchanges, adapting tone, and optimising for continued engagement. When such a system is layered onto a platform used by an Egyptian teenager, the specific cultural and religious vocabulary through which that teenager might naturally express identity, belonging, or distress, is precisely the register in which current Arabic-language AI systems are documented to perform least reliably.

4.3 Family, Religion, and Divergent Mediation Practices

The rights-based literature on children's digital lives has long recognised that mediation, the practices through which parents, schools and communities shape a child's technology use, varies substantially across cultural contexts, and that frameworks developed with Western, individualised family structures in mind do not transfer neatly elsewhere (Livingstone & Third, 2017). Egypt and the wider Arab region are typically characterised by more extended, collective family and community structures, and by a religious framework, predominantly Islamic, that shapes norms around privacy, disclosure and appropriate confidants in ways distinct from the individualised, therapeutic disclosure model implicit in the design of many AI companion features. A chat interface designed, whether by intention or by default, to function as an always-available, non-judgemental confidant sits in some tension with family and community mediation practices in which a child's concerns are more typically expected to pass through, or at least be known to, an extended network of relatives and trusted adults. This is a hypothesis grounded in the existing rights-based and family-mediation literature rather than a claim we test empirically here, but it points to a further, under-examined dimension of double mediation: the second mediation layer is not

only linguistically and culturally skewed in its content, but is frequently built around an implicit model of the child as an autonomous individual user, a model that does not automatically fit every family and community context in which these platforms now operate.

This tension extends to the question of consent and disclosure that runs through the broader children's-rights literature. Stoilova, Livingstone and Nandagiri (2020) show that even in well-resourced Anglophone contexts, children's capacity to meaningfully understand and manage what a platform does with their data depends heavily on literacy support from schools, peers and family. In the Egyptian context, that support structure is itself shaped by an educational system and a set of family expectations organised around collective, rather than purely individual, responsibility for a child's wellbeing, so that the question of who is entitled to know what an AI companion has been told by a child is not merely a technical privacy-settings question but a genuinely cultural one, with no straightforward Western analogue to import wholesale.

4.4 Double Mediation in Practice: Identity Negotiation Online

Bringing these threads together, an Egyptian child or teenager negotiating identity on a mainstream social platform today does so under a distinctive compound pressure. The platform's first mediation already privileges globally engaging, largely Western-coded content through its selection algorithm. Its second mediation, the generative AI layer now built into captioning, translation, moderation and companion features, has been shown in the research literature reviewed above to default toward Western cultural entities and to perform less reliably in Egyptian Arabic than in English or Modern Standard Arabic. The cumulative effect is a platform environment that, for reasons largely invisible to the child using it, structurally nudges identity performance and self-expression toward a cultural register that is not the child's own, precisely the identity-formation concern this Collection's call identifies as central, but here traced to a mechanism, the AI layer, that existing platform-studies scholarship has not yet incorporated into its account of datafied childhood.

It is important to be precise about the nature of this claim. We are not arguing that Egyptian or Arab children are passive recipients of a homogenising cultural pressure; the same rights-based literature this article draws on has consistently shown children to be active, creative negotiators of the platforms and technologies available to them, repurposing tools in locally meaningful ways that designers rarely anticipate. The double mediation framework instead describes the structural conditions under which that negotiation now takes place: conditions that make locally resonant self-expression measurably harder to achieve by default, precisely because the AI systems mediating so much of that expression were not built with an Egyptian teenager's linguistic register or cultural vocabulary as a design priority. Whether, and how, Egyptian children push back against, route around, or creatively reappropriate these defaults is an empirical question this article's conceptual contribution is intended to open up, not to foreclose.

5. Policy and Regulatory Implications

The double mediation framework has direct implications for the current wave of legislative activity aimed at protecting children on social media, activity that this Collection's call explicitly foregrounds. Australia's Online Safety Amendment (Social Media Minimum Age) Act, which took effect in December 2025 and represented the world's first nationwide age restriction of its kind, offers the most closely studied case to date. An early peer-reviewed evaluation published in *The BMJ* in mid-2026, tracking 408 adolescents aged twelve to seventeen before and three months after the law's introduction, found that more than 85 percent of under-16 respondents continued to use restricted platforms at follow-up, most through their own accounts, and that daily use held steady or even rose among some age groups. The study's authors attributed this largely to weak platform-side age verification, most commonly a simple self-declared birthdate, and documented widespread circumvention through borrowed or fake accounts and private browsing. A separate policy briefing reported broadly similar findings from an earlier, independent survey, alongside more mixed public sentiment data suggesting cautious optimism among Australian parents even as the underlying usage data showed limited change.

This early evidence matters for the present argument for a specific reason: the entire architecture of the Australian law, and of the similar proposals now under consideration in Britain, France and elsewhere, targets only the first mediation, platform access itself, through age verification. Even in the counterfactual case where age verification worked perfectly and successfully excluded every under-16 user, the law would have no bearing whatsoever on the second mediation this article describes: the cultural and linguistic defaults embedded in the generative AI systems that legitimately admitted older teenagers, and adults, continue to encounter inside these platforms. An age gate, however well enforced, is silent on whether the AI companion a fifteen-year-old encounters the moment they turn sixteen and gain legitimate access is culturally and linguistically responsive to that child's own context.

Egypt's own emerging legislative response, initiated following President el-Sisi's public call in January 2026 and the subsequent parliamentary statement committing to draft legislation regulating children's social media use, is at an earlier stage and its final shape remains undetermined. The double mediation framework suggests a specific opportunity here: because Egypt's law is still being drafted, it is not yet locked into the access-restriction-only model that dominates the Australian, and likely the British and French, approaches. A regulatory framework informed by the analysis in Section 4 might extend beyond age verification to require, for instance, clear disclosure to child users when they are interacting with generative AI rather than platform content or human interlocutors, a requirement already present in California's companion-chatbot statute and in the EU AI Act's transparency provisions, and might further require platforms operating in the Egyptian market to demonstrate that their AI-generated content and companion features have been evaluated for cultural and dialectal bias in Egyptian Arabic specifically, rather than relying on evaluations conducted primarily in English or Modern Standard Arabic. No major jurisdiction's current or proposed legislation, to our knowledge, includes a requirement of this kind, which we would characterise as a second-mediation-aware regulatory measure, distinct from, and complementary to, the access-restriction measures that currently dominate the policy conversation.

The pattern holds across the other jurisdictions named in this Collection's own call. Britain's government has signalled it will consider a comparable under-16 restriction alongside tighter rules on harmful content and screen time; France's president has pushed for a under-15 ban to be enforceable from the start of a school year; both proposals, like Australia's, are framed principally around access and age verification rather than around what happens inside the platform once access is granted. This is not a criticism of the specific policymakers involved so much as an observation about the current state of the art in platform regulation: the double mediation this article describes is, at the time of writing, effectively invisible to the legislative process in every major jurisdiction we have reviewed, Egypt's own still-forming proposal included.

We note, finally, that the Australian evaluation's own authors caution that longer-term, coordinated assessment is needed before drawing firm conclusions about the ultimate effectiveness of even the first-mediation, access-restriction approach. Our argument is not that age-verification measures are without value, but that they are addressing a necessary, not a sufficient, condition for protecting children's welfare and cultural identity on AI-embedded platforms. A regulatory conversation confined to the question of who may access a platform, without a parallel conversation about what kind of AI-mediated experience awaits those who do, will remain structurally unable to address the identity-formation and cultural-representation concerns this Collection identifies as central to children's and teenagers' relationships with social media.

6. Conclusion and Research Agenda

This article has argued that scholarship on children and social media, and scholarship on children and generative AI, need to be read together rather than as parallel literatures, because the platforms themselves no longer separate the two. We have proposed double mediation as a framework for that integration: a first mediation, the datafied, algorithmically selected environment platform-studies scholarship has long theorised, and a second mediation, the culturally and linguistically uneven generative AI layer now built into that same environment. Applied to Egypt and the wider Arab region, this framework surfaces a distinctive, and so far under-examined, dimension of digital childhood: children whose language, dialect, religious framework and family-mediation practices diverge from the assumptions embedded in dominant AI training data experience a qualitatively different, and by the evidence reviewed here, a less culturally responsive version of the same platforms their peers elsewhere use.

Two limitations of the present analysis should be acknowledged directly. This article is a conceptual and literature-synthesising contribution rather than a report of new primary data on Egyptian or Arab children's platform use, and the double mediation framework, while grounded throughout in published empirical findings on both platform datafication and Arabic-language AI bias, has not itself been tested against direct observation of children's interactions with AI-embedded platforms in this specific regional context. Nor does the framework, as developed here, capture the full diversity within the Arab region itself, whose more than twenty countries vary considerably in connectivity, dialect, religious composition

and regulatory approach; our focus on Egypt is intended as an illustrative, evidence-grounded entry point rather than as a claim of regional uniformity. We regard both limitations as invitations to the empirical research agenda set out below, rather than as objections to the framework itself.

Several lines of future research follow directly from this analysis. First, there is a clear need for empirical, field-based research examining how Egyptian and other Arab children and teenagers actually experience AI-embedded platform features, including AI companions, captioning tools and content moderation, extending the qualitative and survey methods already established in the children's-datafication literature to this newly converged object of study. Second, comparative work across other Global South contexts, South and Southeast Asia, sub-Saharan Africa, Latin America, would test whether the double mediation dynamic documented here for Arabic-language contexts generalises, and how it interacts with each region's own distinct linguistic and family-mediation patterns. Third, as domestically developed Arabic-centric AI systems such as Jais and ArabianGPT mature, research is needed on whether these alternatives meaningfully reduce double mediation in practice, or whether they are adopted too unevenly, and too slowly relative to the dominance of a handful of global platforms, to shift children's everyday platform experience. Finally, we see considerable value in sustained collaboration between childhood and media studies researchers, of the kind this Collection is well positioned to convene, and the computational linguistics and AI-bias research community whose technical findings, reviewed throughout this article, currently circulate largely outside the childhood-studies literature that most needs them. It is only through this kind of interdisciplinary bridge-building that scholarship, and eventually policy, will be equipped to see the whole of what a child now encounters when they open a social media app: not one mediation, but two.

Declarations

Ethics Approval and Consent to Participate

Not applicable. This article is a conceptual and literature-synthesising contribution; it does not report new primary data collected from human participants, and no fieldwork requiring ethics committee review or participant consent was undertaken for this specific manuscript.

Consent for Publication

Not applicable. This manuscript does not contain any individual person's data in any form.

Competing Interests

The authors declare no competing interests.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contribution

N.S. conceived and designed the study, conducted the literature synthesis, and wrote the original draft. N.S. and A.I. contributed to writing – review & editing. F.E. reviewed the manuscript in its final form. N.S. supervised the work. All authors reviewed and approved the final manuscript.

All authors reviewed and approved the final manuscript.

Data Availability

No new datasets were generated or analysed for this article. All demographic, survey and evaluation findings discussed are drawn from the previously published sources cited in the References section, each of which remains available through its original publisher or repository.

References

1. Andoh E (2025), October 1 Many teens are turning to AI chatbots for friendship and emotional support. *Monitor on Psychology*, 56(7). American Psychological Association
2. AraRegBias (2026) Evaluating dialectal and stereotypical bias in Arabic large language models via multi-component metrics. Research Square (preprint)
3. DataReportal (2025) Digital 2025: Egypt. DataReportal / We Are Social / Meltwater
4. Livingstone S, Third A (2017) Children and young people's rights in the digital age: An emerging agenda. *New Media Soc* 19(5):657–670
5. Lupton D, Williamson B (2017) The datafied child: The dataveillance of children and implications for their rights. *New Media Soc* 19(5):780–794
6. Mascheroni G, Siibak A (2021) Datafied childhoods: Data practices and imaginaries in children's lives. Peter Lang
7. Naous T, Ryan MJ, Ritter A, Xu W (2024) Having beer after prayer? Measuring cultural bias in large language models. *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 16366–16393
8. Sallam N et al (2023) Jais and Jais-Chat: Arabic-centric foundation and instruction-tuned open generative large language models. arXiv:2308.16149.
9. Sengar SS et al (2024) ArabianGPT: Native Arabic GPT-based large language model. arXiv:2402.15313.

10. Singer N (2019) *Humanit Social Sci Reviews* 7(5):1321–1337.
<https://doi.org/10.18510/hssr.2019.75171>. The relationship between watching animated cartoon and information processing speed and level for sample children in age group [5–6] years
11. Singer N (2022) Cartoons as the incidental vocabulary acquisition tool for English language learners. *Arab World Engl J* 13(1):330–341. <https://doi.org/10.24093/awej/vol13no1.21>
12. Singer N (2023a) *Positive media psychology: Harnessing the power of media for well-being and resilience*. Eliva Press SRL
13. Singer N (2023b) *The digital generation: Decoding children's media psychology*. Eliva Press SRL
14. Singer N (2024) *Understanding media psychology: Exploring the impact of media on individuals and society*. Eliva Press SRL
15. Singer N (2025) *Reframing and innovating media psychology*. Eliva Press SRL
16. Singer N, Mahmoud E-S, E.-F (2020) Motives of the Egyptian education future for sustainable development: A comparative analysis between 2020 and 2030. *Humanit Social Sci Reviews* 8(2):1 – 28. <https://doi.org/10.18510/hssr.2020.82e01>
17. Stoilova M, Livingstone S, Nandagiri R (2020) Digital by default: Children's capacity to understand and manage online data and privacy. *Media Communication* 8(4):197–207
18. van Dijck J, Poell T, de Waal M (2018) *The platform society: Public values in a connective world*. Oxford University Press
19. Vasan N et al (2025) August 27). Why AI companions and young people can make for a dangerous mix. *Stanford Medicine News Center*
20. Wolfenden L et al (2026) Assessing the early effects of Australia's Social Media Minimum Age Act on adolescent social media use. *BMJ*. <https://doi.org/10.1136/bmj-2026-363695>
21. Associated Press (2026), January 26 Egypt to adopt restrictions on children's social media use to fight 'digital chaos'
22. National Center for Social and Criminological Research (2024) Report on children's exposure to harmful content on social media platforms in Egypt. [Government-affiliated research body, as cited in Associated, 2026]