

Engineering the Mind: A Multidisciplinary Framework for Digital Well-Being

Hame Dhar Dwivedi*

Abstract

Digital technologies now mediate most domains of human activity, yet the engineering disciplines that build these systems have historically treated user well-being as a secondary usability concern rather than a design objective. This paper presents a multidisciplinary review and conceptual synthesis of digital well-being, integrating human-computer interaction, cognitive and motivational psychology, software engineering, and technology law. The objective is to consolidate fragmented, single-disciplined accounts of digital well-being into one coherent framework and to identify the structural gap that has allowed persuasive and manipulative design to persist despite growing scientific and regulatory attention. A narrative-synthesis approach was used, critically appraising peer-reviewed literature drawn from human-computer interaction, psychology, neuroscience, business ethics, and law journals published mainly between 1994 and 2025. The review shows that digital well-being is best explained by three converging strands of evidence: motivational psychology, particularly self-determination theory and its Motivation, Engagement and Thriving in User Experience (METUX) extension, which explains why autonomy-supportive design sustains healthier engagement than restrictive or guilt-based interventions; reinforcement-learning and neurobiological accounts of variable-reward interfaces, which explain compulsive use; and an emerging legal literature that reframes manipulative interface design as a breach of a service provider's duty of loyalty rather than a mere consumer-protection violation. The synthesis further shows that engineering curricula and regulatory instruments remain poorly coordinated with this evidence base. The paper concludes that sustainable digital well-being requires safety-by-design engineering practice, psychologically grounded interface evaluation, and coordinated legal and educational reform, and it proposes actionable recommendations for software engineers, educators, and policymakers.

Keywords: digital well-being; human-computer interaction; self-determination theory; attention economy; dark patterns; technology regulation

INTRODUCTION

The rapid diffusion of information and communication technologies (ICTs) has embedded digital systems into nearly every domain of human activity, from work and education to social relationships and healthcare [1]. This diffusion has produced measurable gains in connectivity, productivity, and access to information, but a growing body of psychological and public-health evidence associates intensive, poorly regulated technology use with reduced subjective well-being, disrupted attention, and compulsive checking behavior, particularly among adolescents [2]. Historically, the software engineering and human-computer interaction (HCI) communities have treated these harms as peripheral usability

*Author for Correspondence

Hame Dhar Dwivedi
E-mail: icameh2024@gmail.com

M. Tech Student, Department of Computer Science and Engineering, Indian Institute of Technology Patna, Bihar, India

Received Date: August 29, 2026
Accepted Date: September 09, 2026
Published Date: September 13, 2026

Citation: Hame Dhar Dwivedi. Engineering the Mind: A Multidisciplinary Framework for Digital Well-Being. Journal of Mobile Computing, Communications & Mobile Networks. 2026; 13(2): 58–68p.

issues rather than first-order design requirements, and the earliest responses to the problem were correspondingly narrow: personal informatics tools such as screen-time counters and app timers that quantify usage without addressing why that usage becomes compulsive [1, 3].

This narrow framing has created a design paradox. Most commercial digital products are optimized for engagement metrics that are only loosely related to user welfare, while the well-being features layered onto these same products are typically restrictive and friction-based, blocking applications, imposing blacklists, or issuing guilt-inducing reminders [3]. Restrictive interventions of this kind treat the user as an object to be constrained rather than a person whose autonomy should be respected, and empirical evaluations show that they are frequently abandoned or produce compensatory overuse once the restriction is lifted [3]. Simultaneously, the underlying interaction design of many platforms continues to rely on variable-ratio reward schedules and other persuasive mechanics that were never intended for therapeutic or restrictive correction in the first place [4, 5]. The result is a structural mismatch: the discipline that creates the stimuli responsible for compulsive engagement is not the discipline that studies the motivational and neurobiological mechanisms those stimuli exploit, nor is it well connected to the legal and educational institutions that could constrain harmful design at scale.

Therefore, the specific problem this study addresses is one of fragmentation rather than the absence of knowledge. Self-determination theory and its user-experience extension, the Motivation, Engagement, and Thriving in User Experience (METUX) model, offer a well-validated psychological account of what makes technology use healthy or unhealthy [6, 7]. Reinforcement-learning and dopaminergic-signaling research explain, at a mechanistic level, why intermittent digital rewards are so effective in capturing attention [4, 5]. Cognitive load theory explains why certain interface patterns overwhelm working memory and degrade decision-making [8]. Distinct legal and business ethics literature has begun to reconceptualize manipulative interface design, commonly termed "dark patterns," as a form of disloyalty by the service provider rather than an isolated consumer protection infraction [9–11]. Each of these literatures has matured largely in isolation, with limited cross-citation between the engineering, psychology, and law communities, and with correspondingly limited translation into either professional engineering practice or computing curricula [1].

This study aims to construct a multidisciplinary framework that connects these separate literatures into a single, coherent account of digital well-being and to use that framework to identify concrete, actionable implications for software engineers, educators, and policymakers. The remainder of this paper is organized as follows. Section 2 reviews the relevant literature across psychology, HCI, neuroscience, and law and identifies the research gap that motivates the present synthesis. Section 3 develops the psychological and behavioral foundations of the framework, centered on the stimulus-organism-response (SOR) paradigm and self-determination theory. Section 4 examines the attention economy through the lens of reinforcement learning and persuasive design. Section 5 addresses cognitive load and interface engineering issues. Section 6 analyzes deceptive design patterns and the legal doctrines emerging to address them. Section 7 compares the global regulatory architectures for digital well-being. Section 8 discusses curriculum reform. Section 9 concludes with a synthesis of the findings and stakeholder-specific recommendations.

LITERATURE REVIEW

Research relevant to digital well-being spans at least four disciplinary traditions, each of which is reviewed here before the cross-cutting gap is identified.

Motivational psychology provides a mature theoretical foundation. Self-determination theory (SDT) holds that human motivation and psychological health depend on the satisfaction of three basic psychological needs: autonomy, competence, and relatedness [7, 12]. When these needs are supported, individuals display more autonomous forms of motivation and greater long-term persistence; when they are frustrated, the result is disengagement, anxiety, or compulsive substitute

behavior [7]. Peters, Calvo, and Ryan extended SDT specifically to digital products through the METUX model, which evaluates a technology's effect on need satisfaction across six nested spheres of experience, from the immediate interface to the surrounding society [6]. Subsequent validation work by Jenö, Diseth, and Grytnes confirmed that interface- and task-level need satisfaction under the METUX model predicts engagement and well-being outcomes in real educational technology settings, lending empirical support to the model beyond its original theoretical proposal [13]. This literature converges on a clear conclusion: autonomy-supportive design, not restrictive design, is a psychologically defensible target for digital well-being interventions. Its main limitation is that it says comparatively little about the specific interface mechanics, such as variable reward schedules, which undermine autonomy in the first place.

This mechanistic gap is addressed by a second, neuroscience-oriented literature. Classic behavioral research on operant conditioning established that variable-ratio reinforcement schedules, in which a reward follows an unpredictable number of responses, produce the most persistent and resistant-to-extinction response patterns of any reinforcement schedule [14]. Schultz's work on dopaminergic signaling subsequently showed that phasic dopamine activity encodes a reward-prediction error rather than the hedonic value of a reward itself, which explains why anticipation of an uncertain reward, rather than the reward itself, drives repeated checking behavior [5]. Lindström and colleagues applied a computational reinforcement-learning model directly to social media posting behavior and found that users' posting latency adjusts dynamically to maximize the net rate of social reward, in a manner formally analogous to reward-learning models used in comparative psychology [4]. This literature is mechanistically precise but normatively silent: it explains how variable-reward interfaces capture attention without specifying what designers or regulators should do about it. Fogg's Behavior Model partially bridges this gap by formalizing the conditions, motivation, ability, and a well-timed prompt, under which any target behavior occurs, and by showing how the same formula that supports beneficial habit formation can be inverted to engineer compulsive use [15, 16].

A third literature, situated in HCI and cognitive psychology, addresses the interface and workload dimensions of digital well-being. Cognitive load theory distinguishes intrinsic, extraneous, and germane load, and shows that poorly designed interfaces impose extraneous load that degrades performance and increases user frustration without any corresponding learning or task benefit [8]. Mark, Gudith, and Klocke's field studies of interrupted work demonstrated that notification-driven task switching increases subjective stress and compensatory work speed even though total task accuracy is preserved, indicating a hidden well-being cost that conventional productivity metrics do not capture [17]. A systematic review and meta-analysis by Roffarello and De Russis of digital self-control tools found that most existing interventions are restrictive rather than autonomy-supportive, and that their effectiveness is inconsistent and often short-lived, reinforcing the SDT-based critique of friction-only design [3]. Population-level survey evidence from Twenge and Campbell further associates higher screen time with lower psychological well-being among adolescents, although the same evidence base cautions that the relationship is moderate in size and moderated by the type, rather than only the quantity, of use [2].

Fourth, more recent literature situates digital well-being within business ethics and law. Mathur et al.'s large-scale audit of shopping websites provided the first systematic taxonomy and prevalence estimate of manipulative "dark pattern" interfaces in commercial practice [9], and subsequent HCI work by Renaud et al. showed that children as young as 11 years can recognize some, but not all, categories of manipulative design, underscoring a design-literacy gap [17]. The United States Federal Trade Commission's own staff report corroborated the rise of increasingly sophisticated manipulative interfaces in consumer-facing products [18]. Bhargava and Velasquez provided a normative business ethics analysis showing that attention-economy business models create structural incentives toward exploitation regardless of individual designers' intentions [11]. Gunawan et al. proposed reframing such practices as "disloyal design," a breach of the service provider's duty to its users, as a doctrinal

alternative to enumerating an ever-expanding list of individual manipulative patterns [10]. In parallel, comparative legal scholarship has documented a growing but fragmented body of technology-specific regulations, including Elvy's analysis of age-appropriate design code mandates in the United States [19], Craven's analysis of the People's Republic of China's "Minor Mode" restrictions on youth device use [20], and Lerouge and Trujillo Pons' comparative study of statutory rights to disconnect from work-related digital communication in France and Spain [21].

Collectively, these four studies show that the scientific and legal knowledge required to design for digital well-being already exists in a mature form within each discipline. What is largely absent is a synthesis that connects psychological need-satisfaction theory, the neurobiological mechanics of persuasive design, cognitive-engineering practice, and the emerging legal doctrine of design loyalty into a single framework usable by software engineers, and a corresponding account of how that framework should be taught. Reviews such as Shin's recent systematic synthesis of human-centered artificial intelligence for digital well-being have begun this integrative work at the level of the SOR paradigm [1], and computing-education guidance such as the CS2023 curricular volume has begun to expand ethics coverage in technical courses [22]. However, neither engineering practice guidance nor curricular reform has yet been explicitly organized around the motivational, neurobiological, and legal evidence reviewed above. This is the gap the present paper addresses: it does not report new primary data, but constructs an explicit cross-disciplinary framework, evaluates it against the reviewed literature, and translates that framework into concrete recommendations for the three stakeholder groups best positioned to act on it: software engineers, educators, and behavioral or legal policymakers.

THEORETICAL FOUNDATIONS: FROM STIMULUS-RESPONSE TO PSYCHOLOGICAL NEED SATISFACTION

The Stimulus-Organism-Response Paradigm in Digital Well-Being

A recurring difficulty in early digital well-being scholarship was the absence of a shared vocabulary connecting system-level design choices and user-level psychological states. Shin's systematic review addresses this by organizing the human-centered artificial intelligence for digital well-being (HCAI-DW) literature around the stimulus-organism-response (SOR) paradigm, an environmental-psychology framework originally developed to explain how environmental cues shape behavior through an intervening internal state [1]. In this adaptation, the Stimulus corresponds to designed and non-designed interaction variables, such as push notifications, algorithmic recommendations, and interface layout; the Organism corresponds to the user's internal cognitive and emotional processing, shaped by both stable traits such as personality and self-regulatory capacity and situational factors such as time pressure; and the Response corresponds to the resulting behavior, attitude, or longer-term psychological outcome [1]. The practical value of this framework is that it redirects engineering attention away from telemetry that only measures the Response, such as total session time, and toward the Organism-level states that mediate whether a given Stimulus produces a healthy or unhealthy Response.

Self-Determination Theory and the METUX Model

Self-determination theory provides the psychological content of the organism component of the SOR paradigm. SDT proposes that human motivation and psychological health depend on the ongoing satisfaction of three basic psychological needs: autonomy, the sense of volition and alignment with one's own values; competence, the experience of mastery and effective action; and relatedness, the sense of connection to others [7, 12]. When these needs are supported, motivation tends toward more autonomous, intrinsically, or identity-congruent forms; when they are frustrated, the theory predicts disengagement, anxiety, and, of particular relevance to technology use, compensatory or compulsive substitute behaviors [7]. Within SDT, organismic integration theory further describes a continuum of motivational regulation, from amotivation through externally controlled regulation to identified, integrated, and finally intrinsic motivation, and shows that

sustained behavior change is best predicted by the more autonomous end of this continuum rather than by externally imposed rewards or guilt [7].

Peters, Calvo, and Ryan operationalized SDT for digital products through the METUX model, which evaluates need satisfaction and frustration across six nested spheres of technology experience: Adoption, the pre-use phase in which expectations about need satisfaction form; Interface, the direct interaction with controls and navigation; Task, the specific functional activity supported; Behavior, the broader habits and routines the technology shapes; Life, the cumulative effect on existential well-being and relationships; and Society, the systemic effect on communities and non-users [6]. Jenö, Diseth, and Grytnes' empirical test of the model in higher-education technology found that need satisfaction at the Interface and Task spheres significantly predicted student engagement, learning, and well-being, providing external validation for the model beyond its originating theoretical paper [13]. The practical significance of the METUX framework for engineering practice is that it makes visible design contradictions that a single-sphere evaluation would miss: a gamified application, for example, may satisfy autonomy and competence needs at the Interface and Task levels through responsive feedback, while simultaneously frustrating autonomy at the life sphere through aggressive notification patterns that displace sleep or offline relationships [6].

THE ATTENTION ECONOMY: REINFORCEMENT LEARNING AND PERSUASIVE DESIGN

Variable-Ratio Reinforcement and Dopaminergic Signalling

The dominant commercial model for consumer software treats user attention as a resource to be captured and monetized, an arrangement that several authors term the "attention economy" [11]. Bhargava and Velasquez argue that this business model creates a structural incentive toward addictive design independent of any individual designer's intent because platform revenue scales with time and attention captured rather than user welfare [11]. The specific mechanism by which attention is captured has deep roots in behavioral psychology. Classic operant-conditioning research established that a variable-ratio reinforcement schedule, in which a reward follows an unpredictable number of responses, produces response patterns that are both highly persistent and unusually resistant to extinction, more so than fixed or predictable reward schedules [14]. Infinite-scroll feeds and notification systems reproduce this schedule: a user who refreshes a feed or checks for a message cannot predict whether that action will yield a socially rewarding outcome, and this unpredictability is itself a reinforcement.

The neurobiological basis of this effect was clarified by Schultz's work on midbrain dopamine neurons, which showed that phasic dopamine release encodes a reward-prediction error, the discrepancy between an expected and an actual outcome, rather than the hedonic magnitude of the reward itself [5]. This finding has an important implication for interface design: dopaminergic signalling drives anticipation and approach behavior ("wanting") even when the eventual reward is unremarkable, which is consistent with the compulsive, low-yield checking behavior widely reported around notifications and social feeds. Lindström et al. formalized this process computationally, applying a reinforcement-learning model to naturalistic social media posting data and showing that individuals adjust their posting frequency to track the moment-to-moment rate of social reward they receive, in a manner directly analogous to reward-learning models developed for non-human animal behavior [4]. Their results indicate that engagement with social platforms is not simply a matter of weak self-control but reflects the operation of a well-characterized, evolutionarily conserved learning mechanism interacting with an interface engineered to maximize reward unpredictability.

The Fogg Behavior Model and Ethical Behavioural Engineering

Fogg's Behavior Model provides an applied design vocabulary that connects these psychological and neurobiological mechanisms to concrete interface decisions. The model states that a target behavior occurs when sufficient motivation and ability converge with an effective prompt at the same

time [15]. Fogg further distinguishes three prompt types: a spark, which adds motivation when ability is already high; a facilitator, which increases ability when motivation is already high; and a signal, which serves only as a cue when both motivation and ability are already sufficient [15, 16]. The same formula that supports beneficial habit formation, such as health tracking or language learning, can be inverted for exploitative ends: by systematically reducing the ability required to perform an action, for example, through single-tap interactions and autoplay, and by algorithmically timing prompts to moments of high user vulnerability, a system can trigger behavior that bypasses considered judgement [15]. The ethical distinction between persuasive and manipulative design in Fogg's framework is not the formula itself but whether the resulting behavior remains congruent with the user's own reflectively endorsed goals, which links this literature back to the autonomy criterion at the center of self-determination theory (SDT) [7, 16].

COGNITIVE LOAD AND HUMAN-CENTRED INTERFACE ENGINEERING

Beyond motivational and reward-based mechanisms, digital well-being is also shaped by the more basic cognitive-engineering question of how much working memory demand an interface imposes. Cognitive load theory distinguishes three categories of mental effort: intrinsic load, arising from the inherent complexity of the task itself; extraneous load, arising from poor presentation, cluttered layout, or inconsistent interaction patterns and wasted on processing the interface rather than the task; and germane load, which is the effortful processing that builds durable understanding and skill [8]. Because working memory capacity is limited, well-designed systems should minimize extraneous load through progressive disclosure, consistent navigation, and predictable interaction patterns, thereby preserving the user's cognitive capacity for germane processing and, ultimately, for the deliberate self-regulation that autonomy-supportive design depends on [8].

Empirical field research on interruptions reinforces the connection between interface design and psychological cost. Mark, Gudith, and Klocke's diary and observational study of knowledge workers found that individuals who were frequently interrupted by notifications completed their work at a faster pace and reported significantly higher stress, frustration, and perceived effort than uninterrupted workers, even though the interruptions did not measurably reduce the accuracy of the completed work [23]. Interfaces that minimize This result is significant for digital well-being because it demonstrates a psychological cost of interface-driven interruption, elevated stress, that is invisible to conventional productivity or engagement metrics and that would not be captured by a well-being intervention narrowly focused on total time-on-device. unnecessary interruption, that batch or defer non-urgent notifications, and that make ongoing background activity transparent to the user address this cost directly and are consistent with the broader cognitive-load and autonomy-supportive design principles established in the theoretical and reward-learning literatures reviewed above.

The practical translation of cognitive load and autonomy principles into working software remains uneven, and a systematic review by Roffarello and De Russis of digital self-control tools found that the majority of deployed interventions still rely on restriction, blocking, or time-limit mechanisms rather than on need-supportive or cognitive-load-reducing design, and that their measured effectiveness is inconsistent across studies and often diminishes over time as users develop workarounds or compensatory usage patterns [3]. This finding is one of the clearest pieces of evidence in the reviewed literature that restrictive design is not merely theoretically suboptimal under self-determination theory but is also empirically underperforming, which strengthens the case for the autonomy-supportive, cognitively considerate design approach advanced throughout this paper.

DECEPTIVE DESIGN PATTERNS AND THE LEGAL RESPONSE

Prevalence and Typology of Dark Patterns

While the previous sections addressed designs that unintentionally undermine well-being, a distinct and more troubling category consists of interfaces deliberately engineered to work against the user's interests. Mathur et al. 's large-scale automated audit of commercial shopping websites provided the

first systematic empirical estimate of how common such "dark patterns" are in practice, documenting recurring categories that include confusing or double-negative "trick wording," undisclosed charges added through "sneaking," asymmetric friction that makes an action easy to start but difficult to reverse ("obstruction," sometimes called a roach motel), emotionally manipulative "confirmshaming" of users who decline an option, and fabricated urgency through false social-proof indicators [9]. These patterns operate by deliberately exploiting well-documented cognitive biases, such as loss aversion and default-option bias, rather than by informing or persuading the user, which places them outside the boundary of ethical persuasive design described by Fogg [9, 15]. Renaud and colleagues' work with 11- to 12-year-old children found that this population could reliably identify some manipulative patterns, such as forced continuity, but consistently failed to recognize more subtle categories such as confirmshaming, indicating that manipulative design imposes a disproportionate burden on users with less-developed critical digital literacy [17]. The United States Federal Trade Commission's own staff report corroborated these academic findings using independent consumer complaint data, concluding that dark patterns had become more prevalent and sophisticated in the years leading up to the report [18].

From Enumerated Patterns to Disloyal Design

The dominant regulatory response to dark patterns has historically been to enumerate specific prohibited patterns within consumer protection statutes, an approach that struggles to keep pace with continually evolving interface techniques and treats each instance of manipulation as an isolated, individually litigated harm rather than as a systemic property of a design relationship [10]. Gunawan et al. proposed an alternative legal framing, "disloyal design," which shifts the locus of analysis from the enumeration of specific interface tricks to the relationship between the service provider and the user: a design is disloyal when an operator exploits its control over interface affordances and the predictable cognitive vulnerabilities of its users for commercial self-dealing that directly harms those users [10]. This reframing is consistent with and doctrinally complements Bhargava and Velasquez's business-ethics argument that the attention-economy revenue model creates structural incentives for exploitation that cannot be fully addressed by regulating individual interface elements in isolation [11]. Read together, this literature suggests that durable protection against manipulative design will require legal standards that address the underlying incentive structure and duty of care owed by a platform to its users, rather than relying solely on a continually expanding list of prohibited interface patterns.

GLOBAL REGULATORY ARCHITECTURES FOR DIGITAL WELL-BEING

Governments in several jurisdictions have begun to translate concerns about manipulative design and compulsive use into binding regulations, and comparing these architectures clarifies both the range of available policy instruments and their respective limitations.

In the United States, Elvy's analysis of age-appropriate design code mandates shows that several states have adopted a "safety-by-design" model requiring platforms likely to be accessed by minors to consider the best interests of child users in their default settings and interface design [19]. Elvy's analysis also documents the constitutional friction this approach has encountered: because such codes regulate design choices that are entangled with expressive content, they have faced First Amendment challenges, and courts reviewing early versions of these statutes have struck down provisions found to be impermissibly vague while upholding more narrowly defined requirements, such as default privacy settings and age-assurance mechanisms [19]. This litigation history illustrates a recurring tension in digital well-being regulation between provisions specific enough to survive constitutional or free-expression scrutiny and provisions general enough to remain adaptable to fast-changing interface design techniques.

China has adopted a substantially more prescriptive approach. Craven's legal analysis of the People's Republic of China's "Minor Mode" regulation documents a framework that requires device manufacturers, application developers, and application stores to jointly implement usage restrictions

for users identified as minors, including curfew-based access restrictions and age-tiered usage limits, and additionally restricts recommendation algorithms from directing certain categories of content to minor users [20]. Craven situates this regulation within a broader international debate about a possible right to Internet access for minors, noting that highly prescriptive, state-mandated restrictions achieve compliance more directly than design-code approaches but raise distinct concerns about parental and individual autonomy and about extension to adult users [20].

The third regulatory model addresses digital well-being not through platform design rules but through labor law. Lerouge and Trujillo Pons' comparative analysis of France and Spain shows that both countries have codified a statutory "right to disconnect" that protects employees from an obligation to respond to work-related digital communication outside contracted hours, with France requiring larger employers to negotiate the specific technical and organizational implementation of this right and Spain requiring implementation in consultation with worker representatives [21]. The authors argue that these national frameworks, grounded in an existing right to privacy in the use of digital devices, could plausibly generalize to a wider European Union standard, though as of their analysis no binding Union-wide instrument had yet been adopted [21]. Together, the design code, prescriptive restriction, and labor rights models address overlapping but distinct dimensions of the same underlying problem. The comparative legal literature suggests that no single regulatory instrument currently addresses the full scope of digital well-being; effective protection is likely to require some combination of design obligations on platforms, use restrictions calibrated to vulnerable populations, and time-based protections situated outside the platform relationship altogether.

EDUCATING THE NEXT GENERATION OF ENGINEERS

A framework for digital well-being has limited practical effect if the engineers who build consumer software are not trained to apply it to their work. The joint Association for Computing Machinery, IEEE Computer Society, and Association for the Advancement of Artificial Intelligence task force responsible for the Computer Science Curricula 2023 (CS2023) guidelines responded to this concern by substantially expanding the Society, Ethics, and Professionalism knowledge area within the recommended undergraduate computing curriculum [22]. Rather than confining ethical reasoning to a single standalone course, the CS2023 guidelines recommend embedding case-based ethical analysis directly within core technical courses such as databases, software engineering, and artificial intelligence, so that students confront questions about data-collection practices, the tension between engagement-maximizing business objectives and user welfare, and the identification of manipulative design patterns as an integral part of learning the underlying technology rather than as a separate humanities requirement [22].

A complementary framework for digital literacy more broadly, rather than computing education specifically, is provided by the European Commission's DigComp 2.2 Digital Competence Framework for Citizens, which, since its most recent update, has included an explicit competence area concerned with protecting health and well-being in digital environments, encompassing the ability to avoid technology-related health risks and to recognize manipulative or addictive design [24]. Read alongside the CS2023 guidelines, this literature indicates an emerging consensus that digital well-being competence, both the capacity to design for it as an engineer and the capacity to recognize threats to it as a citizen, should be integrated into mainstream curricula rather than treated as a specialized elective. However, the present review found comparatively little empirical evidence evaluating the effectiveness of either framework once implemented in practice, a gap that follow-up research should address directly.

CONCLUSION

This study aimed to address the fragmentation of digital well-being scholarship across psychology, neuroscience, human-computer interaction, and law, and to translate a synthesis of that scholarship into guidance usable by the stakeholders best positioned to act on it. Four conclusions can be drawn from this review.

First, the psychological evidence is now reasonably conclusive that autonomy-supportive design, grounded in self-determination theory and operationalized through the METUX model, outperforms restrictive, friction-based interventions as a strategy for supporting healthy technology use, and that this conclusion holds even though restrictive tools remain the dominant form of deployed digital wellbeing feature. Second, the neurobiological and computational reinforcement-learning literature shows that compulsive engagement with variable-reward interfaces is not primarily a failure of individual self-control but the predictable output of a well-understood learning mechanism interacting with interfaces deliberately engineered to maximize reward unpredictability, which shifts a share of the corrective responsibility from the user to the designer. Third, the legal and business-ethics literature is converging on the view that manipulative design is best addressed as a structural, relational problem, a breach of the loyalty a platform owes its users, rather than as a list of individually prohibited interface tricks. This shift has begun to inform comparative regulation in the United States, China, and the European Union, each of which has adopted a materially different regulatory instrument for a closely related underlying harm. Fourth, engineering and digital-literacy education is beginning to incorporate this evidence base through initiatives such as the CS2023 curricular guidelines and the DigComp 2.2 framework, but this integration remains recent, unevenly adopted, and not yet supported by strong outcome evidence.

The principal contribution of this paper is the integration itself: by connecting motivational psychology, reinforcement-learning neuroscience, cognitive-engineering practice, and technology law within a single explanatory account organized around the stimulus–organism–response paradigm, the paper offers software engineers, educators, and policymakers a shared vocabulary for a problem that has historically been addressed by each discipline in isolation. Three sets of actionable implications follow directly from this synthesis:

For software engineers and system architects, the evidence supports replacing purely restrictive well-being features with adaptive designs that support user self-awareness and goal-behavior alignment, minimize extraneous cognitive load through consistent and predictable interfaces, and eliminate deceptive patterns by ensuring that opt-out and cancellation processes are no more difficult than opt-in and subscription processes.

For educators, particularly in computing and information technology programs, the evidence supports embedding case-based ethical analysis directly into technical courses following the CS2023 model and aligning digital-literacy instruction with recognized competence frameworks such as DigComp 2.2 so that students learn to recognize, and not only to build, manipulative design.

For psychologists, behavioral researchers, and policymakers, the evidence supports designing clinical and regulatory interventions around the satisfaction of autonomy, competence, and relatedness rather than around external restriction or guilt, and pursuing coordinated legal instruments, spanning platform design obligations, use restrictions for vulnerable populations, and time-based protections such as the right to disconnect, since no single regulatory model reviewed in this paper addresses the full scope of the problem.

This review is subject to limitations inherent to a narrative multidisciplinary synthesis rather than a systematic review with pre-registered search criteria: source selection, while guided by disciplinary breadth and the methodological standing of the cited work, cannot claim the exhaustiveness of a formal systematic review, and rapidly evolving regulatory developments, particularly ongoing constitutional litigation over design-code statutes, may outpace the legal analysis summarized here. Future work should pursue systematic empirical evaluation of autonomy-supportive digital well-being interventions at scale, longitudinal outcome studies of the CS2023 and DigComp 2.2 educational reforms, and comparative empirical assessment of the design-code, prescriptive-restriction, and labor-rights regulatory models identified in this review. Notwithstanding these limitations, the convergence

of independent evidence from psychology, neuroscience, engineering, and law toward a common conclusion, that sustainable digital well-being requires designing with, rather than against, human autonomy, provides a reasonably firm foundation for the practical recommendations advanced above.

Acknowledgments

The author gratefully acknowledges the Department of Computer Science and Engineering, Indian Institute of Technology Patna for the institutional support during the preparation of this manuscript.

REFERENCES

1. Shin Y. Toward human-centered artificial intelligence for users' digital well-being: systematic review, synthesis, and future directions. *JMIR Hum Factors*. 2025;12(1).
2. Twenge JM, Campbell WK. Associations between screen time and lower psychological well-being among children and adolescents: evidence from a population-based study. *Prev Med Rep*. 2018;12:271-83.
3. Roffarello AM, De Russis L. Achieving digital wellbeing through digital self-control tools: a systematic review and meta-analysis. *ACM Trans Comput Hum Interact*. 2023;30(4):1-66.
4. Lindström B, Bellander M, Schultner DT, Chang A, Tobler PN, Amodio DM. A computational reward learning account of social media engagement. *Nat Commun*. 2021;12(1):1311.
5. Schultz W. Predictive reward signal of dopamine neurons. *J Neurophysiol*. 1998;80(1):1-27.
6. Peters D, Calvo RA, Ryan RM. Designing for motivation, engagement and wellbeing in digital experience. *Front Psychol*. 2018;9:300159.
7. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol*. 2000;55(1):68-78.
8. Sweller J. Cognitive load theory, learning difficulty, and instructional design. *Learn Instr*. 1994;4(4):295-312.
9. Mathur A, Acar G, Friedman MJ, Lucherini E, Mayer J, Chetty M, et al. Dark patterns at scale: findings from a crawl of 11K shopping websites. *Proc ACM Hum Comput Interact*. 2019;3(CSCW):1-32.
10. Gunawan J, Hartzog W, Richards N, Choffnes D, Wilson C. Dark patterns as disloyal design. *Ind LJ*. 2024;100:1389.
11. Gunawan J, Hartzog W, Richards N, Choffnes D, Wilson C. Dark patterns as disloyal design. *Ind LJ*. 2024;100:1389.
12. Deci EL, Ryan RM. *Intrinsic motivation and self-determination in human behavior*. New York: Springer Science & Business Media; 2013.
13. Jenö LM, Diseth Å, Grytnes JA. Testing the METUX model in higher education: interface and task need-satisfaction predict engagement, learning, and well-being. *Front Psychol*. 2021;12:631564.
14. Skinner BF. *The behavior of organisms: an experimental analysis*. Cambridge (MA): B.F. Skinner Foundation; 2019.
15. Fogg BJ. A behavior model for persuasive design. In: *Proceedings of the 4th International Conference on Persuasive Technology*; 2009 Apr 26. p. 1-7.
16. Fogg BJ. Persuasive technology: using computers to change what we think and do. *Ubiquity*. 2002;2002(December):2.
17. Nelson V. Computing curriculum computer engineering (CCCE): a model for computer engineering curricula in the next decade. In: *2004 Annual Conference*; 2004 Jun 20. p. 9-332.
18. Mark G, Gudith D, Klocke U. The cost of interrupted work: more speed and stress. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*; 2008 Apr 6. p. 107-10.
19. Renaud K, Sengul C, Coopamootoo K, Clift B, Taylor J, Springett M, et al. We're not that gullible! Revealing dark pattern mental models of 11-12-year-old Scottish children. *ACM Trans Comput Hum Interact*. 2024;31(3):1-41.
20. Zac A, Huang YC, Von Moltke A, Decker C, Ezrachi A. Dark patterns and consumer vulnerability. *Behav Public Policy*. 2025:1-50.

21. Elvy SA. Age-appropriate design code mandates. *U Pa J Int Law*. 2023;45:953.
22. Craven T. Kids, no phones at the dinner table: analyzing the People's Republic of China's proposed "minor mode" regulation and an international right to the Internet. *Chi J Int Law*. 2024;25:219.
23. Lerouge L, Trujillo Pons F. Contribution to the study on the 'right to disconnect' from work: are France and Spain examples for other countries and EU law? *Eur Labour Law J*. 2022;13(3):450-65.
24. Pérez-Escoda A, Fernández-Villavicencio NG. Digital competence in use: from DigComp 1 to DigComp 2. In: *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality*; 2016 Nov 2. p. 619-24.