



# Cultivating *sīla* Online: the use of Cognitive Interventions in Systems Design

**Paul Trafford**

This is an unpublished conference paper for the 9<sup>th</sup> Annual Jubilee Centre for Character and Virtues conference at Oriel College, Oxford University, Friday 10<sup>th</sup> – Sunday 12<sup>th</sup> September 2021.

These papers are works in progress and should not be cited without author's prior permission.

Jubilee Centre for Character and Virtues

University of Birmingham, Edgbaston, Birmingham, B15 2TT United Kingdom

T: +44 (0) 121 414 3602 F: +44 (0) 121 414 4875

E: [jubileecentre@contacts.bham.ac.uk](mailto:jubileecentre@contacts.bham.ac.uk) W: [www.jubileecentre.ac.uk](http://www.jubileecentre.ac.uk)



## **Abstract**

This paper addresses from a Buddhist perspective the burgeoning challenges of living and flourishing with digital technology. At its heart are the principles of human agency and a framework for the cultivation of *sīla* (moral virtue). These allow us respectively to develop criteria for evaluating new technology, such as artificial intelligence, and to design systems that enhance well-being and human relationships. Accordingly, we apply them to the design of novel online social networks, where we treat the problem of maintaining heedfulness by deploying thinking routines to strengthen cognitive (and hence moral) functioning.

## 1. Introduction

The Covid-19 pandemic with its faltering responses has exacerbated many issues that were already looming large before the crisis began. There has been a major impact on mental health, arising especially from the increased interaction with and absorption in the digital world at the expense of physical contact. Today, just a few technology companies provide global online services that typically occupy our mental landscape hours a day, every day. The widespread distress and discontent that has arisen in relation to such services is often attributed to (what feels like) their incessant demands, a response that is characteristic of *dukkha*, a Buddhist term in Pali for deep-seated dis-ease.

One response to the burgeoning challenges of such over-exposure is to ration or avoid – hence the banning of mobile phones in schools and around the dinner table with colleagues. Whilst this gives a breathing space, we are still faced with their calls for attention when we have to use these digital services. So, another response is to better accommodate them through instruction in how they work and in how to properly configure the preferences according to one's needs. More sophisticated training can take fuller account of the dynamics of communication in these environments, such as inculcating a critical digital literacy for online social interactions (Tagg *et al* 2019).

In these computer-mediated interactions, software can also offer assistance to human adaptation, particularly through what are technically termed *interventions*, the introduction of specific measures to treat an identified need. Such interventions have been used in particular to deal with mental health issues, ranging from the regular use of clinically-approved meditation apps for stress (NHS 2021) to more sophisticated environments, including gaming worlds that train adolescents with cognitive behavioural therapy (Merry *et al* 2012, Fleming *et al* 2021). Even before the pandemic

started, there was a general increase in such issues, though their severity appears to be increasing (Hafstad and Augustin 2021). This has prompted significant investment in developing novel, evidence-based projects to tackle the situation, such as training adolescents to develop resilience in the digital world (UKRI 2021).

Whilst these interventions have been shown to have positive impact by increasing resilience and emotional self-regulation, they do not address directly the source of the problems – the design of the software services themselves. For social media, especially, despite the much-publicised furor with the Cambridge Analytica scandal and the heightened concern around data privacy (Venturini and Rogers 2019, Véliz 2021), many other serious issues are hardly mentioned. There remains a general acceptance, perhaps resignation, that these services, though deeply flawed, will not fundamentally change; it is ‘normal’ and will remain so in the ‘new normal’. Being adept at navigating legal impositions, such services may feel free to continue largely as before and even ramp up their operations. Yet, just as a physical environment affects one’s social development, so too an online environment affects one’s cognitive, emotional and spiritual development; a digital tsunami can be overwhelming no matter how resilient one may be.

There is therefore an urgent need to pay close attention to the design of these systems and their architecture, to better understand their effects on individuals and society and, then, to make fundamental improvements. As has been highlighted, particularly among moral philosophers, such a situation poses unique ethical challenges and requires a response that can properly incorporate education in character and moral virtue, in ways that are most appropriate to online contexts, whether that be through technomoral virtues (Vallor 2016) or cyber-wisdom (Dennis and Harrison 2021). Our orientation is practical, as we seek to show how virtue-based approaches can play a key role at critical points in software systems design to aid greater well-being. Giving particular attention

to social networking sites ('SNS' for short), we propose a framework for human relationships based on Buddhist ethics. This new approach provides insights into systems designs that can remedy current systemic flaws and aid human flourishing.

We then explore how digital environments impact the mind, finding affinities with current research in cognitive science. By highlighting cognitive aspects, we are able to indicate how interventions are already intrinsic to systems design and precisely targeted, but in a detrimental way, as a core means of subconsciously fostering addictiveness. By turning this around, we propose interventions to increase awareness and restore conscious control to the user. Educational interventions to address moral concerns have already been called for by Dennis and Harrison (as above). The interventions proposed here offer a response based on Buddhist notions of virtue and new kinds of SNS architecture, which are designed to enhance human agency and strengthen virtuous conduct.

## ***2. Moral Conduct and Well-being***

Even in the early years of the Web, when the Internet was a relatively simple environment, it was recognised that a moral perspective was needed (Hambridge 1995):

In general, rules of common courtesy for interaction with people should be in force for any situation and on the Internet it's doubly important where, for example, body language and tone of voice must be inferred.

Whilst there was misbehaviour in cyberspace almost from its inception, most of it was on a small scale and due to particular users, not the software (Sternberg 2012). In contrast, as Internet usage has mushroomed, previous affordances have been turned into liabilities through poor systems

design, with evidence of major harmful effects even for the moderators (Oireachtas 2021). Whilst legislative bodies are slowly being stirred into high-level responses to particular issues (DCMS 2020: section 8), moral considerations remain to be treated more thoroughly by all parties involved in the software lifecycle, particularly in the early stages of requirements analysis and systems architecture. To bring these issues to the fore, we therefore seek answers to the question: How do we support, nurture and safeguard especially wholesome and fruitful personal relationships in systems design?

### *Sīla*

The importance of virtuous behaviour in daily life is evident in the following observation from the Dhammapada, a popular anthology from the Buddhist canon:

All mental phenomena have mind as their forerunner;  
mind is their chief, and they are mind-made. If one  
speaks or acts with a defiled mind, then suffering  
follows one even as the wheel follows the hoof of  
the draught ox.

All mental phenomena have mind as their forerunner;  
mind is their chief, and they are mind-made. If  
one speaks or acts with a pure mind, then happiness  
follows even as one's shadow never leaves.

(Dhp 1,2)<sup>1</sup>

---

<sup>1</sup> The sutta reference style follows mainstream conventions for English translations, as used on the Access to Insight website, <https://www.accesstoinsight.org/abbrev.html>. Similarly, where names of translators are given, their translations are to be found on this site. All other translations are the author's, usually based on others.

Thus, everyone's happiness is primarily the responsibility of each individual, where the moral quality of mind determines whether or not the result is conducive to well-being. Moral conduct is termed *sīla* (to give the Pali spelling) and its recommended cultivation is succinctly expressed as: 'Cease to do evil. Cultivate good. Purify the mind.' (Dhp 183).

The practice may be more fully expressed as a mental discipline and disposition of wholesomeness that leads to an internal sense of ease. It applies to all actions in body, speech and mind that sow the seeds of well-being. Such actions, termed *kamma* in Pali (*karma* in Sanskrit), are intentional and it is intention which is the main driver:

"Intention, I tell you, is *kamma*. Intending, one does *kamma* by way of body, speech, & intellect." (AN 6.63)

Historically, a code of conduct, termed the *Vinaya*, was introduced to support the *sangha* community, those who had left the householder life to pursue the training. For lay Buddhists, it has become customary to observe a reduced and simplified code called 'the Five Precepts'. These are usually formulated as *vāriṭṭa* (avoidances), namely: refraining from taking life; from taking that which is not given; from sexual immorality; from false speech; and from intoxicants that lead to heedlessness. Observing these rules generally prevents problems and leads to greater harmony in society, with more energy for making productive contributions. This is easier to see when one considers the *cāriṭṭa* (positive performance) counterpart, which is to cultivate compassion, generosity, contentment, honesty and heedfulness – qualities widely regarded as the characteristics of a model member of society.

In social sciences, the main concept underpinning the evaluation of well-being in social media has been *social capital*. The analysis of human activity is generally in terms of social relationships, with the theory and supporting methods tending to emphasize sociability as a good in itself. As

notions of social capital have been extended, they have been increasingly oriented around notions of ‘civicness’, despite warnings that such extensions carry risks, particularly of equating relationships with resources (Portes 98).

In contrast, the Buddhist perspective remains rooted in the individual and their motivations, with *sīla* being fundamental to actions both individual and in relationships, so any evaluation of social relationships should take this into consideration. The individual’s desirable stock or worth to humanity and society is most properly encapsulated by the notion of *puñña*, which is translated as ‘merit’ or ‘good deeds’<sup>2</sup>. Its opposite is *apuñña* (or, synonymously, *pāpa*), de-merit or ‘bad deeds’, which are what we might refer to today as ‘toxicity’ or undesirable stock. *Puñña* is constituted by three kinds of karma; the first two we have already mentioned, namely, *dāna* (generosity) and *sīla*, with the third being *bhāvanā* (mind development), being practices in concentration and mindfulness (AN 8.36). This kind of ‘capital’ has the distinct feature of contributing to society’s wholesomeness or purification.

### **Self-regulation**

Perhaps Buddhism’s greatest impact on mainstream health has been the introduction of mindfulness practice in clinical care (NHS 2018), including Wellbeing apps for NHS Staff (NHS 2021). It is a form of *self-regulation*, which is a term used especially in clinical parlance concerning the control of one’s behaviour, usually tied to specific long-term goals. The training in awareness of the present moment is widely taught and practised as an effective means to treat stress, depression and other clinical conditions. However, in traditionally Buddhist countries, it is properly situated in religious practice and is invariably connected to *sīla*, whereby the clear comprehension that arises through

---

<sup>2</sup> For an introduction to *puñña* and its role in cultivating the spiritual life see Thanissaro Bhikkhu, *Merit: A Study Guide*, Access to Insight (BCBS edition), 2005 (revised 2013) <https://www.accesstoinsight.org/lib/study/merit.html>



mindfulness arising from a practice of purity. Mindfulness therefore goes hand in hand with heedfulness:

Ever grows the glory of him who is energetic, mindful and pure in conduct, discerning and self-controlled, righteous and heedful.

(Dhp 24, Acharya Buddharakkita trans.)

This kind of mastery over oneself involves controlled behaviour in regards to the inputs through the sense doors (sight, hearing, touch, taste, tactile sensations, and mental objects). Generally, the restraint applies to actions in body, speech and mind. One particular application of self-regulation, that it most pertinent to social media, is expressed in the Vācā Sutta on ‘five-star’ speech.

"It is spoken at the right time. It is spoken in truth. It is spoken affectionately. It is spoken beneficially. It is spoken with a mind of good-will." (AN 5.198)

Applying this to status updates, we find:

<b>Vācā Sutta</b>	<b>Status updates in Facebook, Twitter etc.</b>
Spoken at the right time	Spoken at any time (with nudges)
Spoken in truth	May or may not be spoken in truth
Spoken affectionately	May be spoken with or without affection
Spoken beneficially	May or may not be beneficial

Spoken with a mind of good-will	May or may not be spoken with a mind of good-will
---------------------------------	---

**Table 1: Comparative behaviours between virtue-based and social media in practice**

This in itself already highlights how the affordances in social media are liable to issues such as rampant hate speech (‘hateful speech’ is more correct from a Buddhist perspective because it conveys intention) and the phenomenon of Twitter storms.

### **Friendship**

What does friendship really mean and how best to nurture it? Largely for the sake of efficiency, SNS have generally persisted with deploying a single connection type, labelling it ‘friend’. Some so-called ‘friends’ are acquaintances, and some are even unknown to the user. In this scenario, describing relationships is reduced merely to ‘friends’, ‘friends of friends’, ‘close friends’, and so on, resulting in a flattened landscape. This effectively dilutes relationships to their lowest common denominator, a problem that was identified early in Facebook’s history (Stein 2007) and which provided the stimulus for the author’s initial investigations into alternative architectures (Trafford 2007). The use of privacy options, including the addition of arbitrary groups, have provided some finer measure of control, but subsequent revisions by Zuckerberg (2010) have not adequately tackled the problem, as evidenced by increased polarisation (Tagg *et al* 2017). Moreover, whereas nurturing friendship tends to take time, SNS connections can be made instantly, prompting further questions about the value of such ‘friendship’.

From a Buddhist viewpoint, the right kind of friendship is essential to well-being in society. The Buddha used the term *kalyāṇamittatā*, which means ‘association with good friends or good friendship,’ and defined it thus:

Herein, Vyagghapajja, in whatsoever village or market town a householder dwells, he associates, converses, engages in discussions with householders or householders’ sons, whether young and highly cultured or old and highly cultured, full of faith, full of virtue, full of charity and full of wisdom. He acts in accordance with the faith of the faithful, with the virtue of the virtuous, with the charity of the charitable and with the wisdom of the wise.

(AN 8.54, Narada trans.)

Cultivating this kind of friendship underpins a trusted network. We next consider how that network may be visualized.

### **The Sigālovāda Sutta and the Six Directions**

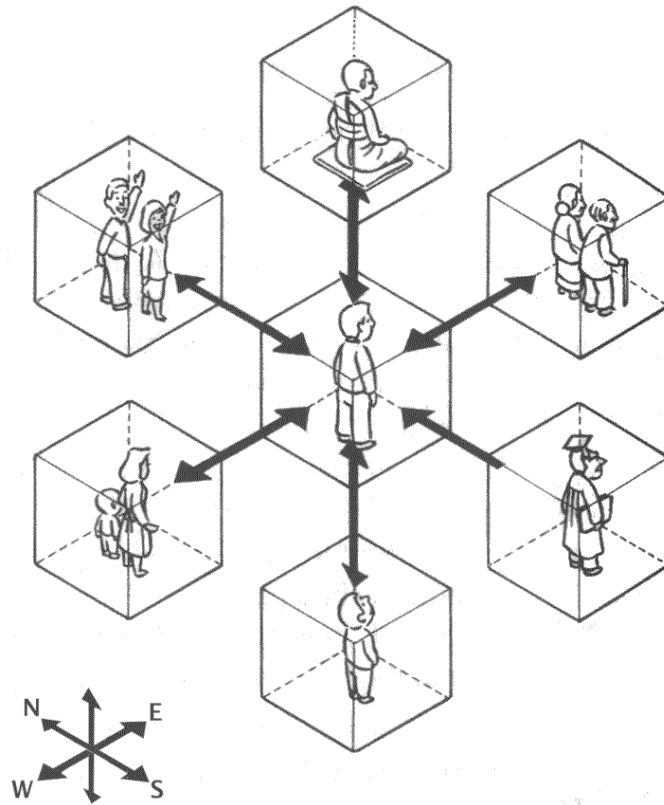
In the Sigālovāda Sutta (DN 31), one of the few teachings he gave on the household life, the Buddha drew on his familiarity with domestic situations to demonstrate how, in particular, kinship relations serve as the basis of social relations. He speaks at length on meritorious behaviour and the cultivation of true friendship, teaching Sigāla, a young householder, how to properly nurture human relationships. He uses visualisation, structuring his teachings through a multidimensional model spanning six quarters or directions – the four cardinal points, plus above and below. Each direction corresponds to a type of relationship: parents to the East, spouse and children to the West, teachers to the South, friends and associates to the North, servants and employees to the nadir, ascetics and brahmans to the zenith.

Examining the model more closely, there are six kinds of dyadic relationship, with the reciprocity rooted in a sense of moral responsibility and sincere concern for the welfare of others:

1. spiritual guide or guru – disciple
2. teacher – student
3. friend or associate (peer relationship)
4. employer – employee
5. spouse (peer relationship)
6. parent – child

The formulation should also be understood in the cultural context; as a householder, Sigāla was privileged – much of his daily business was conducted at home, with many of his employees being servants and workpeople – and activities outside the home were conducted along the caste lines established in Indian society at the time. Today, this situation has changed somewhat and there are numerous cultural variations across the world. Hence, we would not expect to replicate this model exactly. Even so, most of the architecture is universally true and a system design can take account of local variation and communicate this to users as part of the process of mutual cultural enrichment.

A visual depiction using cubes has been provided by Venerable Dattajeevo (2006), representative of Thai society (Figure 1):



**Figure 1: The Six Directions of the Sigalovāda Sutta  
(after Dattajeevo)**

This 3D representation is strongly suggestive of a tessellation to fill the entire space, indicating how local ties lead to global interconnectedness. Furthermore, these ties are orthogonal – severing connections in one direction leaves intact ties in the other directions because of the separation of concerns.

This immediately suggests a radical change in viable frameworks for new SNS designs, in which the single ‘friendship’ connection is replaced by a pre-defined choice of relationship types, to be mutually agreed, thereby articulating different social ties (Trafford 2011). Further, the sutta provides the foundations for a suitable privacy architecture (what to share with whom) along with

guidance for appropriate behaviour (according to the relationship type). In terms of the linguistics underlying communication, it also provides a more nuanced handling of the problem of *addressivity* so that one can be more at ease to share personal matters, in contrast to the common coping strategy – among those who take care – of avoiding or limiting sensitive communication in existing SNS (Tagg and Sergeant 2019).

Relationships evolve gradually and organically in any number of ways in the various online social contexts. This too can be incorporate in systems design by restricting the functionality to the degree to which connections know each other (as mutually agreed). According to Buddhist teachings, as long as the actions are wholesome, the relationships will flourish.

### **3. Agency and Cognitive Processes**

Thus far, we have argued for the moral imperative and provided a social context for practice. We now focus on the cognitive dimension and start with another quote of the Buddha:

Household life is crowded and dusty; life gone forth is wide open. It is not easy, while living in a home, to lead the holy life utterly perfect and pure as a polished shell.

(MN 36, Ñāṇamoli & Bodhi trans)

The main difficulty for householders is the host of incumbent duties and responsibilities, all of which have psychological overheads – things to become attached to, to worry about, etc. –the mind is seldom allowed to settle. If life has always been mentally ‘crowded and dusty’, then life online increases the cognitive burden even more, echoed in daily complaints of ‘information overload’. With regards to the practice of moral virtue, neuroscientists have already alerted us an underlying

problem – moral decision-making takes time (Immordino-Yang et al 2009, Marziali 2009). If technology is robbing us of our ability to function at such a basic level, then we need new methods that fundamentally protect who we are.

### **Human agency and criteria for evaluating technology**

Human *agency* is typically defined as the exercising of one's ability to act<sup>3</sup>. Here we adopt additionally the sense of having the ability without necessarily exercising it. As outlined previously, the Buddha stated that agency is manifest through *karma* and is intentional. To more fully convey what this means, he gave a short and relatively little-known teaching, the Attakāri Sutta (The Self-Doer) (AN 6.38), expressing agency as a doer (oneself or other) having six characteristics: initiation, exertion, making effort, steadfastness, persistence and endeavouring.

To illustrate, we choose a couple of familiar examples: an electronic calculator and satellite navigation ('satnav' for short). On the one hand, a calculator can increase autonomy by helping us to carry out more complex sums to fulfil various tasks at home and at the office. It can generate exact answers with almost no effort, in just a few taps. Similarly, satnav largely removes the burden of route planning on roads, detailing every stage of the journey. On the other hand, a reliance on a calculator can weaken our ability to reason in everyday situations; sometimes we have to make a snap decision, but with no time to press buttons, we miscalculate. There are also many scenarios when we don't need exact numbers, only estimates, which we may derive through visual observation – to ascertain how congested a place is, whether there are enough boxes to transport a collection of books, and so on.

---

<sup>3</sup> As with, for example, the definition of 'agency' provided in the Stanford Encyclopaedia of Philosophy, <https://plato.stanford.edu/entries/agency/>

Practical drawbacks already suggest use of technology should be judicious, but there is a growing body of neurological research that shows a reliance on digital aids comes at the expense of regions in the brain dedicated to such tasks. It is now well-documented how the ability to carry out arithmetic and personal navigation are impaired, with possibly far-reaching impacts on learning (Baron 2021). As we lose this ability, there are practical ramifications; for instance, we increase the risk of accepting incorrect output after accidentally typing in a wrong number because we have lost a more general sense of quantity.

The importance is magnified for technology that affects mental agency. This is most acutely the case for Artificial Intelligence (AI), which broadly covers software that emulates human intelligence – or otherwise seemingly engages in thought processes – through various methods such as natural language processing and machine learning.<sup>4</sup> It has potentially critical impact as it starts to become involved with decision-making and important life choices. The complexity and sophistication of such systems make for an altogether greater task for the evaluator; ensuring that AI-enhanced environments are healthier requires understanding not only the interplay between the human mind vis-à-vis machines and the algorithms being generated, but also the wider implications.

A detailed discussion of AI's ramification lies outside the scope of this paper, but with reference to the Attakāri Sutta we can, at least, draw attention to areas that need particular consideration. Even with elementary examples of technology, impacting other limbs of agency has wider knock-on effects. For example, exertion and steadfastness are general traits of character that typically require a degree of concentration and habitual activity. Always reaching for a labour-saving gadget or some frictionless facility reduces perseverance and is more likely to foster a habit of laziness. Not persisting leads to avoidance of difficult problems and of the pursuit of their solution,

---

<sup>4</sup> A comparison of human intelligence with machine intelligence and its significance for human agency depends on how creativity is regarded. This is discussed in relation to mathematics in (Trafford 2021: Chapter 3).



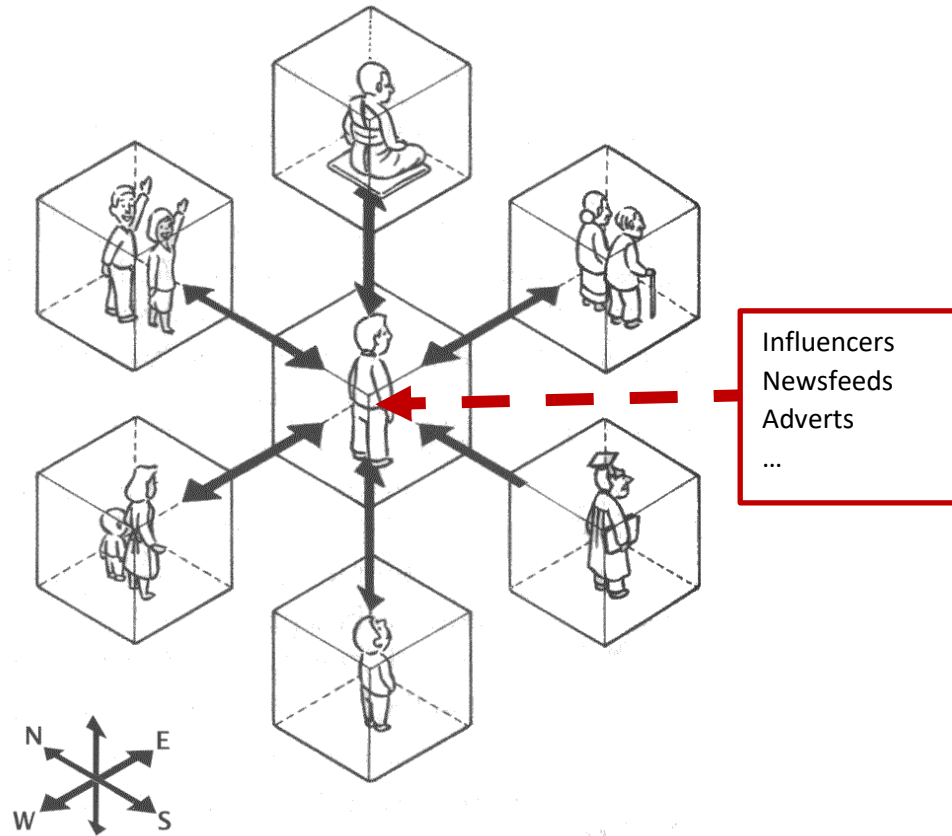
removing the opportunity to gain fresh insights. For society, habitually making an effort is essential for maintaining healthy human relations and for reaching out and forming new ones.

The following sections will focus on social media to articulate some of the issues, with particular reference to algorithms, which are the ‘thinking routines’ that underpin software, whether or not they involve AI. Understanding the techniques behind their use paves the way for possible solutions.

### **Information flows**

The values that we hold dear are increasingly influenced and even determined by online activities; younger generations especially use social media to find out about current affairs in preference to traditional news media. With the exception of direct messaging, the user interface of such systems aggregates a huge volume of content irrespective of how much has been self-generated. In Facebook, the timeline displays items from newsfeeds, stories and third-party advertisements. There are options to control what is displayed according to permissions (based on public, friends or selected individuals) or at a very granular level, per content block (such as pages). However, it is a fiddly exercise and does not generally prevent various personal and impersonal content being mixed together. Furthermore, reactions are displayed from friends of friends, so that much of what’s displayed is not coming from direct contacts.

If we map this to our 6-dimensional model, we see that the information architecture of the timeline substantially bypasses all the standard dyadic connections (Figure 2):



**Figure 2: Digital Bypass of The Six Directions**

The streams of content in the timeline would be much clearer if instead we were to apply a set of filters reflecting the ‘Six Directions’ architecture, displaying only news and stories contributed only by family, friends, teachers, and so on. Contacts could, as now, mention products as items of interest or recommendations, where the provenance would be more clearly sign-posted. The display of third-party advertisements could become an opt-in preference to reduce their intrusiveness. Whilst businesses may prefer opt-out, it should be noted that recent changes have increasingly prioritised the more meaningful interactions between contacts (Mosseri 2018, Sethuraman *et al* 2019).

## Addictive user interface design

Modern life is characterised by an increase in consumer choice, where little thought is required to make selections. However, decision-making concerning more durable aspects of life, such as relationships, requires deliberate and considered reflection. This is in jeopardy as online engagement increases in intensity, with a multitude of distractions from personal gadgets and social media services.

A significant factor is *user interface design* (now more commonly referred to as *user experience design*). One of the most influential works is Steve Krug's 'Don't make me think' (2000), extolling the principle of user interfaces that require minimal brain power to figure out; everything should be self-evident, so each page should be uncluttered, laid out in a sound and orderly way; navigation through the site should be straightforward and intuitive. This 'no brainer' appears to be common sense, but the push for ease of use assumes a habitual free-flowing mode of activity with no pause for thought, justified by claims of users having so little time. The principle also contrasts with much of actual user experience – of pop-ups for cookie selections, of embedded advertising, and of elements jumping about as the page loads, and so on. Whilst claiming that all users are different, designers make major assumptions about what they have in common and what they need. With search engines being ubiquitous, instructions manuals no longer accompany electronic goods, but they are often missed and there is a reduction in literacy about the goods (Pogue 2017).

Of greater concern is the incorporation of this principle in the design of social media. It is a key instrument supporting the goal of 'frictionless sharing' extolled by Mark Zuckerberg. In the technology world, 'friction' correlates with effort, but sharing on social media should be effortless. This is reflected in information flows: the timeline uses infinite scroll and when one story video ends the next one automatically starts playing. Yet, the lack of friction also correlates with a lack of

deliberation – especially on what’s appropriate to share. Without friction, it’s all too easy to share accidentally, especially when other design principles such as ‘nudging’ are there to help us along.

In recent years, a few veterans of the tech industry have created a stir with frank remarks about social media: in 2017, Sean Parker, former Facebook Director, confessed that Facebook had deliberately incorporated elements designed to foster addiction through a series of ‘dopamine hits’ (Allen 2017). Shortly after, Jaron Lanier (2018) was exhorting everybody to ditch their social media accounts because of their harmful addictiveness. However, this is not a new realization. Reporting on his Trinidadian fieldwork in (2010), Daniel Miller observed contrasting aspects of ‘Fas’ book; whilst being a facility to coordinate life-saving responses to disasters, it also had a tendency to aggravate a less desirable side of Trinidadian culture, exacerbating “a national characteristic leading to the disorder of bacchanal”. We see the “affordances”, which have been commonly extolled since (Ellison *et al* 2007), encouraging instant gratification and a lack of mindfulness, precipitating breakdowns in human relationships.

A more sober assessment has come from psychiatrists, who have moved to suggest that social network addiction be added to the spectrum of Internet spectrum disorders. (Karaiskos *et al* 2010) Many design elements permeate from other disciplines, particularly gaming, where the frictionless principle features strongly, particularly in terms of reducing physical resistance to the input controls (as exemplified by a well-lubricated Rubik’s Cube). With advances in processing power and network bandwidth, software and hardware environments have become more sensorially realistic. The combined ease and sophistication have become the building blocks of addictiveness; and gaming finds itself included in the WHO categorisation of addictions alongside Internet addiction (WHO 2015, WHO 2021).

#### 4. Interventions

Parker indicated that various devices from behavioural psychology are deployed to keep user engaged or locked in. A fundamental technique is the use of interventions. Etymologically, *intervention* is derived from two Latin words: *inter*, which means ‘between’, and *venire*, meaning ‘to come’. Thus, an intervention is an act of coming between someone on an existing course or path and the continuation of that path. That ‘act’ could be a natural phenomenon, as in ‘rain stopped play’, but, more usually, there is some intention behind the intervention with a view to modifying the outcome.

With systems, interventions can be designed on behalf of the platform maker to foster increased habitual use and encourage less thoughtfulness. They can also be designed on behalf of the user to do the opposite, that is, to reduce habitual use and increase thoughtfulness. We shall discuss both below.

##### **Attention and interventions**

The ‘attention economy’ has become a means for large organisations to view attention as a scarce economic resource. Applied to Internet use, systems are designed to retain and nurture our attention and steer or nudge us in particular directions as consumers. Social media companies are adept at minimizing relevant information shared with services outside of their control, whilst cramming in a great deal of information within their service walls and promising more. For example, the developers generally include an email notification function with their services to provide timely updates such as comments from friends. They typically design their interventions to provide only numerical summaries or one-line headlines about a friend’s update, usually without excerpts of what the friends have written. The user is thus obliged to click through and sign in to the service to find out further information.

Once logged in to the platform, everything is geared to retaining the attention. Interventions or, rather, micro-interventions, can be embedded at any point in the user journey – such as a web page pop-up prompting us to sign up to a newsletter. They can simply display a few additional words of text in a strategic placed based on the user’s profile and activity, or indicate tiny changes such as incrementing the number of ‘likes’; Parker provides a frank and alarming summary concerning Facebook:

“The thought process that went into building these applications ... was all about: 'How do we consume as much of your time and conscious attention as possible?’” “And that means that we need to sort of give you a little dopamine hit every once in a while, because someone liked or commented on a photo or a post or whatever. And that's going to get you to contribute more content, and that's going to get you ... more likes and comments.” “It's a social-validation feedback loop ... exactly the kind of thing that a hacker like myself would come up with, because you're exploiting a vulnerability in human psychology.” (Allen 2017)

This strategy is so invasive that it is debatable whether resilience measures alone can be effective. It is conceivable that companies would even sponsor such initiatives if it means they can continue with their methods. Indeed, in an ironic twist, the habitual and addictive patterns of behaviour encouraged by these platforms have also enabled a class of healthy interventions. Social media have been used successfully for campaigns targeting other kinds of addictions such as smoking and alcohol abuse (Naslund *et al* 2017, Bonar *et al* 2020). Notwithstanding these successes, the addictiveness of these systems should still be addressed and can be done so in a way that still supports existing positive uses.

One effective response that has always been available, as offered in the Axios interview, is to cut off the cord and not use social media. This continues to have its advocates, who argue how liberating it is (Newport 2019). However, far more so now than in the past, it will significantly limit and in many cases cut off communication lines with relatives, friends and colleagues.

Particularly astute observations about the design of the systems have come from experienced industry insiders, who understand the issues through both broad technical knowledge of the platforms and practical experience. However, most of those employed in the industry have kept fairly quiet (or else have not been reported) about interventions, perhaps because they are, by nature, included without the users' express permission. It tends to be left to commentators to communicate the insights. Thus, suggestions have already been made concerning the user interface, particularly to strengthen cognitive abilities. For example, Amber Case makes astute observations on how SNS designs have been generally optimised for speed, frequency and continuity of engagement, thereby removed opportunities to pause for thought, to think before acting, leading to many ruptures (2018). She then goes on to make suggestions – in her section, 'Designing for Warmer Engagement' – of system interventions for 'cooling off'. These include introducing a delay before comments get out of hand on a contentious thread, prompting a user to reflect on a topic before commenting on it, and a pause on new additions to feeds.

Case is mindful that these kinds of changes for a more humane experience face considerable inertia. Some momentum might be gathered by trawling the Web for such suggestions and creating a knowledge base to inform future designs, but concrete proposals will need to emerge in order to have a greater chance of having a significant impact.

## Thinking Routines

Social media connects people in abstract digitally-mediated settings that are far removed from face-to-face communication. As frictionless user interfaces tend to discourage thought, additional means are required to strengthen thinking processes that encourage and support reflection on what is appropriate behaviour and on how to take care in initiating and developing social relationships.

To enhance this capacity, we outline a methodology concerning *thinking routines*, derived from a large body of educational theory and practice developed by Harvard University's Project Zero, whose history goes back to the 1960s (Gardner 2016). These routines are designed to nurture a culture of thinking, to enhance a person's capacity to enquire and learn, and to instil this as a habit. Inspiration is drawn from the Innovating with Intelligence project, particularly around character and intellect, which developed materials "that help teachers foster student's thinking dispositions through the exploration of six thinking ideals: truth, beauty, imagination, understanding, fairness and self-direction." (Ritchhart 2002: xxx) The project has gradually developed a compendium of Thinking Routines: for example, the routine, 'See/Think/Wonder' facilitates gaining new knowledge.<sup>5</sup> Visible Thinking is an extension designed for the classroom to make thinking apparent to enable teachers to assist students more effectively in their thought processes (Ritchhart *et al* 2011).<sup>6</sup> We consider it instructive for wider application as it helps to authenticate knowledge, which in the online social context, is a vital consideration when people have been reduced to avatars and multimedia profiles.

As the application domain concerns the social sphere, we propose new thinking routines specific to the cultivation of wholesome relationships. Consistent with the Buddhist view of agency

---

<sup>5</sup> *Thinking Routines Toolbox*, Project Zero, Harvard Graduate School of Education.  
<http://www.pz.harvard.edu/thinking-routines>

<sup>6</sup> *Visible Thinking*, Project Zero, Harvard Graduate School of Education.  
<http://www.pz.harvard.edu/projects/visible-thinking>



expressed in the first verses of the Dhammapada (as quoted above), we argue for the exploration of a further ideal: moral virtue, with the language for routines extended accordingly. Here the instantiation uses *sīla*, but the approach can be applied similarly to other ethical codes of conduct. We note in particular that whilst the processes encourage open enquiry, they should also assist in decision-making – about people, not concepts or inanimate objects – according to an ethical code of conduct.

There are various ways such thinking routines can be implemented. For example, they can be recommended in a way similar to annual information security training commonly used in corporate environments. We can apply these routines in controlled environments as a form of scaffolding for those in the early stages of learning how to use the Internet and social media. The main audience is likely to comprise mainly young children, the elderly and those with special needs. Outside such environments other kinds of provision can be made; whilst options might be more limited for make the thinking visible, options for imagining and visualization remain available.

In line with preceding discussion, to encourage these habits of enquiry among the general population, we suggest that cognitive interventions be incorporated in the design of social media services, applied to specific contexts. Technically, this can be implemented directly in the social media platforms themselves or via third-party apps or plugins. To illustrate, we focus on two common activities, in which we present new routines for making a connection and sharing a status update. In each case the routines break down the considerations into steps and a series of prompts or questions.

The first use case is to apply 'Five Star Speech' for status updates, encouraging users to pause and think before sharing. The five statements lead directly to questions to be presented for what

might be labelled, 'Mindful Thinking Routines' or the 'Thinking Routine for Five Star Speech'.

**1. Notice**

What is it that you 'see' that prompts you to write? What's inspiring you? How does it make you feel? Is it positive or negative?

**2. Think**

What do you want to say?

Is it true? Have you got your facts right?

Is it the right time to say it?

**3. Imagine**

Who's going to read your message?

What will they feel when they read it?

How will it benefit them?

**4. Plan**

What's the best way to say it? Prompts:

- friendly intentions

- gentleness

(Why do these work?)

On submission, the system can validate the input using machine learning and check particularly for offensive language. If something is flagged as potentially inappropriate, then a further intervention can come into play:

Did you really mean to say that? It may be regarded as offensive ... Please try to rephrase.

The second use case concerns invitations to connect. Here our thinking routine reflects the teachings to Sigāla:

### **1. Identify**

Who are they? How do I know them? Are they who they say they are?

### **2. Think**

How well do I know them?

Do I want to associate with them online?

This part could be facilitated by a further Thinking Routine based on other passages in the sutta. For example, with respect to friends, the sutta describes how one should minister to friends and associates (as the North). This can prompt one to ask oneself about what one is ready to give (rather than take).

### **... Consider**

Am I ready to:

(i) be generous?

- (ii) show courtesy in speech?
- (iii) offer help?
- (iv) be impartial?
- (v) act with sincerity?

In turn, consider: is the associate a true friend?

...**Imagine** you are in the company of this person.

Do they act as follows?

- (i) Protect me from being heedless?
- (ii) Protect my property if I am heedless?
- (iii) Are they someone I can turn to when in danger?
- (iv) Are they someone who sticks around when I'm in trouble?
- (v) Do they show consideration for my family?

### **3. Do they fit?**

Weigh up the evidence for and against making the connection.

How do you feel about them? What's your intuition?

The guidance might conclude that if all is fine, then now is the time to send the invitation. If writing a note of introduction, then the Thinking Routine for Five Star Speech can be invoked.

Considering the impact of the intervention on speech, we observe that the process of initiation is mindful and no longer a stimulus response, which is an increase in autonomy. The words are not the first thing to enter the head, but arise through a deliberate and careful process that

requires exertion and effort. To sustain this approach will require each of the qualities of steadfastness, persistence and endeavouring. Altogether, agency is enhanced and this approach can be applied to any existing system.

For the second intervention, the evaluation of the suitability of a connection – in terms of authentication and one’s own commitment – foreshadows many of the characteristics of agency as a kind of mental roleplay. However, with the current architectures based on the single connection type, in many cases insufficient criteria will be satisfied, with the result that requests are turned down, so the opportunity of nurturing long-term friendships may be curtailed. On the other hand, through separation of concerns, as depicted in the teachings to Sigaala, a more appropriate level of expectation is set and hence there is greater likelihood to establish and develop a wholesome relationship.

## **5. Conclusions**

The framework and methods proposed in this paper are rooted in *sīla*, the Buddhist notion of moral virtue. They offer a constructive response to the pressing ethical issues around the design and use of digital technology, especially social media, aiming to compensate for the absence of many cultural norms in online communications and, thus, to restore greater respect and consideration.

Underpinning the approach is the observation that moral decision-making depends on sufficient cognitive headroom, motivating the use of thinking routines as interventions. The proposals are preliminary sketches and need further work, refinement and testing, but should give sufficient indication of their value in creating and protecting the mental space that allows reflection on *sīla* in one’s actions and interactions. It is hoped that this sufficiently intimates a new and safer direction of travel in the design and development of such online services.

Taking all these observations on board, we propose a process that reviews any existing or future technological systems in terms of their impact on human agency, considering in turn each characteristic of the Attakāri Sutta. The focus should be on designs and cover the entire technology lifecycle from concept to service. When surveying systems, critical junctures should be determined and assessed for their impact on agency – against the six criteria – in the short and long term. Systems may thus be designed in ways that intrinsically support the cultivation of *sīla*, effectively building in ‘self-care’ for oneself and others without reliance on specialist apps.

Whilst indicative of the potential for enhancing human flourishing, such measures will only gain traction through the development of practical demonstrations built from the ground up according to these principles. Accordingly, there is a need to develop of a prototype SNS, to act as a kind of maquette to test out these proposals, to explore and refine different configurations. Some initial steps are described on the author’s project site (<https://research.siga.la/>), but such an initiative needs to be scaled up and become a collaborative and interdisciplinary effort involving many parties. Our aspiration is, in any case, to a future in which advanced technology is used wisely to support a wholesome society imbued with *sukha* (happiness) rather than *dukkha*.

## References

- Allen, M. (2017). ‘Sean Parker unloads on Facebook: “God only knows what it’s doing to our children’s brains”’. Axios, 9 November 2017, <<https://www.axios.com/sean-parker-unloads-on-facebook-god-only-knows-what-its-doing-to-our-childrens-brains-1513306792-f855e7b4-4e99-4d60-8d51-2775559c2671.html>>
- Baron, N. S. (2021). ‘Know what? How digital technologies undermine learning and remembering’, *Journal of Pragmatics*, 175: 27–37. DOI: 10.1016/j.pragma.2021.01.011

- Bonar, E. E., Schneeberger, D. M., Bourque, C., Bauermeister, J. A., Young, S. D., Blow, F. C., Cunningham, R. M., *et al.* (2020). 'Social Media Interventions for Risky Drinking Among Adolescents and Emerging Adults: Protocol for a Randomized Controlled Trial', *JMIR Research Protocols*, 9/5: e16688. DOI: 10.2196/16688
- Case, A. (2018). 'The problems with Facebook are inherent in its design, but that can change'. *TechCrunch*, 16 April 2018, <<https://social.techcrunch.com/2018/04/16/the-problems-with-facebook-are-inherent-in-its-design-but-that-can-change/>>
- Dattajeevo, V. (2005). *Man's Personal Transformation*. (A. Methmanus, Tran.). Bangkok: Dhammakaya Foundation.
- Dennis, M., & Harrison, T. (2021). 'Unique ethical challenges for the 21st century: Online technology and virtue education', *Journal of Moral Education*, 50/3: 251–66. Routledge. DOI: 10.1080/03057240.2020.1781071
- Department for Digital, Culture, Media & Sport (2020), *Online Harms White Paper*. Department for Digital, Culture, Media & Sport, Home Office.  
<<https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper>>
- Ellison, N. B., Steinfield, C., & Lampe, C. (2007). 'The Benefits of Facebook "Friends:" Social Capital and College Students' Use of Online Social Network Sites', *Journal of Computer-Mediated Communication*, 12/4: 1143–68. DOI: 10.1111/j.1083-6101.2007.00367.x
- Fleming, T., Lucassen, M., Stasiak, K., Sutcliffe, K., & Merry, S. (2021). 'Technology Matters: SPARX – computerised cognitive behavioural therapy for adolescent depression in a game format', *Child and Adolescent Mental Health*, 26/1: 92–4. DOI: 10.1111/camh.12444

Gardner, H. (2016). *Harvard Project Zero: A Personal History* | Project Zero.

<<http://www.pz.harvard.edu/resources/harvard-project-zero-a-personal-history>>

Hafstad, G. S., & Augusti, E.-M. (2021). ‘A lost generation? COVID-19 and adolescent mental health’, *The Lancet Psychiatry*, 8/8: 640–1. Elsevier. DOI: 10.1016/S2215-0366(21)00179-6

Hambridge, S. (1995). ‘Request For Comments: 1855: Netiquette Guidelines’. Internet Engineering Task Force. <<https://tools.ietf.org/html/rfc1855>>

Immordino-Yang, M. H., McColl, A., Damasio, H., & Damasio, A. (2009). ‘Neural correlates of admiration and compassion’, *Proceedings of the National Academy of Science*, 106/19: 8021–6. DOI: 10.1073/pnas.0810363106

Karaiskos, D., Tzavellas, E., Balta, G., & Paparrigopoulos, T. (2010). ‘P02-232 - Social network addiction : a new clinical disorder?’. *European Psychiatry*, Vol. 25, p. 855. Presented at the 18th European Congress of Psychiatry, March 27, Munich, Germany. DOI: 10.1016/S0924-9338(10)70846-4

Krug, S. (2000). *Don't Make Me Think: A Common Sense Approach to Web Usability*. New Riders.

Lanier, J. (2018). *Ten Arguments for Deleting Your Social Media Accounts Right Now*. New York: Henry Holt & Company.

Marziali, C. (2009). ‘Nobler Instincts Take Time - USC News’. *USC News*. 14 April, 2009.

<<https://news.usc.edu/29206/Nobler-Instincts-Take-Time/>>

Merry, S. N., Stasiak, K., Shepherd, M., Frampton, C., Fleming, T., & Lucassen, M. F. G. (2012).

‘The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial’, *BMJ*, 344: e2598. British Medical Journal Publishing Group. DOI: 10.1136/bmj.e2598



- Miller, D. (2010). ‘Fas’ book (Facebook) in Trinidad’. *Material World*, 7 February 2010.  
<<http://www.materialworldblog.com/2010/02/fas-book-facebook-in-trinidad/>>
- Mosseri, A. (2018). ‘Bringing People Closer Together’. *About Facebook: News*, 11 January 2018.  
<<https://about.fb.com/news/2018/01/news-feed-fyi-bringing-people-closer-together/>>
- Naslund, J. A., Kim, S. J., Aschbrenner, K. A., McCulloch, L. J., Brunette, M. F., Dallery, J., Bartels, S. J., *et al.* (2017). ‘Systematic review of social media interventions for smoking cessation’, *Addictive Behaviors*, 73: 81–93. DOI: 10.1016/j.addbeh.2017.05.002
- Newport, C. (2019). *Digital Minimalism: Choosing a Focused Life in a Noisy World*. Portfolio.
- NHS. (2018). ‘Mindfulness’. *nhs.uk*. <<https://www.nhs.uk/mental-health/self-help/tips-and-support/mindfulness/>>
- . (2021). ‘NHS England » Wellbeing apps’. <<https://www.england.nhs.uk/supporting-our-nhs-people/support-now/wellbeing-apps/>>
- Oireachtas, House of the. (2021). ‘Joint Committee on Enterprise, Trade and Employment debate – Wednesday, 12 May 2021’. text.  
<[https://www.oireachtas.ie/en/debates/debate/joint\\_committee\\_on\\_enterprise\\_trade\\_and\\_employment/2021-05-12/2](https://www.oireachtas.ie/en/debates/debate/joint_committee_on_enterprise_trade_and_employment/2021-05-12/2)>
- Pogue, D. (2017). ‘What Happened to User Manuals?’. *Scientific American*, 316/4: 30. DOI: 10.1038/scientificamerican0417-30
- Portes, A. (1998). ‘Social Capital: Its Origins and Applications in Modern Sociology’, *Annual Review of Sociology*, 24: 1–24.
- Ritchhart, R. (2002). *Intellectual Character: What It Is, Why It Matters, and How to Get It*. John Wiley & Sons.

- Ritchhart, R., Church, M., & Morrison, K. (2011). *Making Thinking Visible: How to Promote Engagement, Understanding, and Independence for All Learners* | Wiley. Jossey-Bass (Wiley).
- Sethuraman, R., Vallmitjana, J., & Levin, J. (2019). 'Using Surveys to Make News Feed More Personal'. *About Facebook*, 16 May 2019. <<https://about.fb.com/news/2019/05/more-personalized-experiences/>>
- Stein, J. (2007). 'You Are Not My Friend'. *Time*. 4 October, 2007.  
<<http://content.time.com/time/subscriber/article/0,33009,1668441,00.html>>
- Sternberg, J. (2012). *Misbehavior in Cyber Places: The Regulation of Online Conduct in Virtual Communities on the Internet*. University Press of America.
- Tagg, C., & Seargeant, P. (2019). 'Context design and critical language/media awareness: Implications for a social digital literacies education', *Linguistics and Education*, Digital language practices: media, awareness, pedagogy, 62: 100776. DOI: 10.1016/j.linged.2019.100776
- Tagg, C., Seargeant, P., & Brown, A. A. (2017). *Taking Offence on Social Media: Conviviality and Communication on Facebook*. Palgrave Macmillan.
- Trafford, P. (2007). 'On "Friends" and other associations | EDUCAUSE'. *Paul's blog*. EDUCAUSE Blogs, 10 November 2007.  
<<https://web.archive.org/web/20120608142212/http://www.educause.edu/blog/pault/OnFriendsandotherassociations/167285>>
- . (2011). 'Supporting *Kalyāṇamittatā* Online: New Architectures for Sustainable Social Networking'. Presented at the The 3rd World Conference on Buddhism and Science (WCBS), December 1, Mahidol University, Thailand.
- . (2021). *Buddhism and computing*. Oxford: Mud Pie Books.

- UKRI. (2021). ‘£24 million investment into adolescent mental health’. *UKRI News*, 28 June 2021.  
<<https://www.ukri.org/news/24-million-investment-into-adolescent-mental-health/>>
- Vallor, S. (2016). *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting*. New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780190498511.001.0001
- Véliz, C. (2021). ‘Privacy is Power’. Penguin.
- Venturini, T., & Rogers, R. (2019). ““API-Based Research” or How can Digital Sociology and Journalism Studies Learn from the Facebook and Cambridge Analytica Data Breach’, *Digital Journalism*, 7/4: 532–40. Routledge. DOI: 10.1080/21670811.2019.1591927
- World Health Organization. (2015). *Public health implications of excessive use of the internet, computers, smartphones and similar electronic devices: meeting report, Main Meeting Hall, Foundation for Promotion of Cancer Research, National Cancer Research Centre, Tokyo, Japan, 27-29 August 2014*. World Health Organization.
- . (2021). *ICD-11 - ICD-11 for Mortality and Morbidity Statistics: 6C51: Gaming disorder* (Classification No. ICD-11). World Health Organization.  
<<https://www.who.int/standards/classifications/classification-of-diseases>>
- Zuckerberg, M. (2010). ‘Mark Zuckerberg - From Facebook, answering privacy concerns with new settings’. *Washington Post*, 24 May 2010. <<http://www.washingtonpost.com/wp-dyn/content/article/2010/05/23/AR2010052303828.html>>

(Websites retrieved on 31 August 2021.)