

# WEDF Student Initiative Submission

## Social Media Design, Autonomy and Snudges

*Kyle van Oosterum*

### **Introduction**

It has become popular in digital ethics to critique the ways in which digital technologies and social media hijack attention, strain ordinary human communication and compromise their users' autonomy (Vallor, 2012; Véliz, 2019; Williams, 2018). Of these three concerns, the notion of autonomy has also figured predominantly in the history of ethical and political thought as well as in contemporary philosophical thinking. Autonomy is often construed as the crucial ability to control various aspects of one's life in accordance with one's conception of the good (Mill, 2010[1859]). Concerns about the adversarial design of these technologies and how they infringe upon individual autonomy have been well-stated and warrant creative thinking about potential long-term solutions. While important theoretical work is being done to think about re-designing social media in ethical ways, such widespread changes are unlikely to occur in the near future. This warrants a distinct approach, one which though unfamiliar, is no less important.

This distinct approach focuses on what social media users *themselves* can do, at this moment, to regain their autonomy in spite of the adversarial design of digital technologies. Short of boycotting their use – a highly impractical solution – it is clear that users must learn to re-orient aspects of these technologies so that it serves *their* purposes, not Twitter or Facebook's bottom line. In interviews about Thaler & Sunstein's (2008) *Nudge*, they have both made playful references to what is called a 'snudge' or self-nudge. This is a highly effective way for individuals to impose limits on themselves for their own good. One example of a snudge might be keeping your phone in another room while you are attending a Zoom meeting to prevent phone-related distractions. Much like how Odysseus had to tie himself to the mast of

his ship to resist the temptations of the sirens, social media users could impose restrictions on themselves to use digital technologies in a manner that suits and respects their autonomy.

To defend the nudging approach, a brief outline will first be conducted to explain the ways in which social media design and digital technologies are adversarial. Afterward, concrete examples of nudges will be suggested that may bolster the autonomy of digital technology users along with an empirical justification of their effectiveness. Finally, concerns about the nudging approach will be addressed to nevertheless conclude the valuable role of self-regulation in a challenging digital environment.

### **Social Media as Adversary**

Though users of social media often experience pleasure by indulging in engaging content, scrolling through their news feed and being virtually connected to others, it is no secret that these media are designed to *keep* one's attention with little intention of letting it go (Rosas, 2012; Williams, 2018). Williams (2018) has called this design feature of digital technologies 'adversarial' as it manipulates the user into making decisions they may not have wanted to make. There are many features of social media design that accomplish this goal, but it will suffice to name two to understand just how adversarial such technology can be.

#### *Eyal's Hook Model*

Human beings are particularly good at forming habits and this habit-forming tendency has been hijacked by social media design. Eyal's (2014) 'Hook Model' is a simple four-step process by which many social media networks exploit habit-formation to get users unthinkingly hooked. First, a *trigger* – something which gets individuals to engage, e.g.: a loud ping or a red-notification bubble above an app-icon. This exploits an evolutionarily formed tendency to respond to new stimuli, regardless of what the stimulus is. Second, if the trigger is successful it prompts an *action*, which often involves clicking, tapping or engaging with the trigger. This leads to the key step in the Hook Model, the delivery of a *variable reward*, that is, a surprising advertisement, video or image. This starts to wire the brain to produce the neurotransmitter dopamine in anticipation of the variable reward. Finally, this gets individuals to undertake *investment*, putting in their own time, data or money. This provides social media designers with analytics to improve the efficacy of triggers and strengthen the overall Hook Model.

It is no accident that Eyal calls it a ‘hook’ model, as it operates using the same principles that get people addicted to problematic behaviors such as gambling and drug-taking. It causes users to spend too much doing things that either distract them from their other goals or make them feel some level of regret. On the simplest understanding of autonomy, it should be clear that this element of social media design is disrespectful to a user’s capacity to make decisions that reflect what they value as important for a good life.

### *Devious Defaults*

Another element of this adversarial design is what may be termed *devious defaults*. These are the non-incidental built-in settings or features that guide users to invest their time and money into social media websites. Devious defaults come in two varieties: *fixed defaults* and *variable defaults*. Fixed defaults are built-in settings or features that cannot be altered or require great technological savvy to change. Examples of fixed defaults include the red-notification bubbles that appear on app-icons. Neuroscientists and psychologists have documented how red colors capture attention more strongly than other colors (Kuniecki et al., 2015). Notably, the color of these notification-bubbles cannot be changed anywhere in a device’s settings. They can only be turned off, which is highly impractical. Many users perceive notifications to have useful information for them or remind them to connect with others. The choice here is made deliberately unfair: notification or no information.

More commonly, variable defaults are those built-in settings or features that can be changed. One example of a variable default is a toggle-switch to turn ‘auto-play’ off on YouTube. This gives users the possibility to switch off the feature that makes a new video play around five seconds after the current video has ended. While variable defaults are far more helpful than fixed defaults, it is well-known that users rarely alter their default settings at all, let alone on their technological devices (Thaler & Sunstein, 2008). For example, one study showed that 95% of users would not change the configuration of their Microsoft Word settings (Spool, 2011). This percentage is likely to be the same, if not higher, given the attention-grabbing nature of these technologies. In short, devious defaults will either make it impossible to change design settings or depend upon a user’s laziness to not alter them at all. This challenges a user’s ability to control the decisions they make and will often lead them to become more vulnerable to the Hook Model.

### **Snudging, Precommitment and Autonomy**

The above section describes just two ways in which social media design can be adversarial toward users; these ways are far from exhaustive. Assuming that a sudden change in human motivational psychology is unlikely, users of social media will have to level the playing field to resist adversarial digital technologies. This is where the snudge or self-nudge comes into play. Snudging is essentially a form of *precommitment*. Precommitment strategies consist of voluntarily imposing restrictions on oneself to promote one's goals and prevent distraction and susceptibility to external stimuli. One example of precommitment could be giving your wallet to a friend on a night out to prevent yourself from buying cigarettes. A wealth of psychological studies have shown precommitment to be widely effective in helping individuals make better choices in line with their long-term goals, values and preferences. (Studer et al., 2019). Though the Hook Model exploits a human tendency for instant, effort-free gratification, precommitment strategies aim to attack this model on two levels: either it (1) removes the 'trigger' or (2) discourages the action with a punishment (e.g.: paying your friend a fine if you bought cigarettes).

What are some examples of nudges in the context of social media use? 'StayFocusd' is an extension that can be added to a web-browser which allows users to set a time-limit for whichever website prevents them from staying focused on a certain task. It is a snudge as it imposes a limit on users to help resist the temptations of indulging in social media. Once a user reaches their allotted time-limit, the website becomes blocked which removes the trigger and stops the Hook Model in its tracks. Another popular type of snudging tool is the app 'Beeminder' which allows users to put in their credit card and agree to get charged \$5 if they fail to meet a certain goal they have set (e.g.: spend less time on Facebook). By consistently reminding them to enter their data every day, as well as knowing that users want to avoid a 'sting', this weakens the Hook Model by discouraging certain actions with a punishment.

At this point one might object that nudges are merely 'productivity boosters' and that they are not really connected to autonomy. An example of a productivity booster could be the 'Pomodoro' method, a time-management technique wherein one breaks down work into 25-minute intervals punctuated by 5-minute breaks. Here, it seems that a productivity booster is tied to accomplishing relatively short-term tasks, ones that might not be so meaningful in a person's life. One might seek to resist this definition, arguing that the difference between a productivity booster and a snudge may depend on the *attitude* one takes towards a given method or technique. Indeed, the heterogeneity of values and goals that individuals have

seems to suggest that, in principle, the distinction between a snudge and a productivity booster is tough to pin down.

Notwithstanding, while a particular snudge can boost productivity, this objection ignores the fact that precommitment strategies are very much tied to helping individuals achieve meaningful long-term goals that form part of what constitutes a good life. On this understanding of nudges, they are uniquely connected to autonomy in providing the *means* by which individuals actually control and carry out whichever *end* they wish to pursue. Far and above a productivity booster, which seeks to improve the execution of one's short-term tasks, nudges offer individuals a powerful weapon to fight against adversarial digital technologies. Like the original concept of the *nudge* as developed by Thaler and Sunstein (2008), the snudge operates under the assumption that human beings are fallible and often need guidance to achieve what they want. At the moment, the designers of digital technologies are attempting to deliver what they think users want; the snudge ensures that, for now, users will decide for themselves what they *know* they want.

## **Conclusion**

While the focus on nudges for users to uphold their autonomy is valuable, it should be promoted alongside the need for ethically reforming social media design. It is clear that social media networks wield a large amount of power over what people think, feel and ultimately do with a large portion of their lives. Digital ethical codes must be implemented that reign in such power to redesign social media to be our partner and not our adversary (Véliz, 2019). That being said, waiting on institutional reform is neither sufficient nor necessary for individuals to assert their digital autonomy now. Nudges do not simply boost one's productivity, they are the means by which fallible human beings can nevertheless align themselves with goals worth wanting. At no point should individuals be presented with a choice to either use digital technologies and accept adversarial design, or impractically disconnect oneself from the virtual world. It is consistent with enjoying the benefits digital technologies provide that we can insist on fixing what is wrong with them. Rather than make the question about whether or not to use social media networks, the focus should be on how to design them so that we can use them better. For now, social media users can get by with a simple snudge in the right direction.

## **References**

- Eyal, N. (2014). *Hooked: How To Build Habit-Forming Products*. Penguin.
- Kuniecki, M., Pilarczyk, J., & Wichary, S. (2015). The color red attracts attention in an emotional context. An ERP study. *Frontiers in Human Neuroscience*, 9(APR), 1–14. <https://doi.org/10.3389/fnhum.2015.00212>
- Mill, J. S. (n.d.). *On Liberty*. Penguin.
- Rosas, O. (2012). Types of Internet Use, Well-Being, and the Good Life. In P. Brey, A. Briggie, & E. Spence (Eds.), *The Good Life in a Technological Age* (pp. 215–223). Routledge.
- Spool, J. (2011). *Do users change their settings?* UIE Archive. <https://archive.uie.com/brainsparks/2011/09/14/do-users-change-their-settings/>
- Studer, B., Koch, C., Knecht, S., & Kalenscher, T. (2019). Conquering the inner couch potato: precommitment is an effective strategy to enhance motivation for effortful actions. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 374(1766), 20180131. <https://doi.org/10.1098/rstb.2018.0131>
- Thaler, R., & Sunstein, C. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press.
- Vallor, S. (2012). New Social Media and the Virtues. In P. Brey, A. Briggie, & E. Spence (Eds.), *The Good Life in a Technological Age* (pp. 193–202). Routledge. <https://doi.org/10.4324/9780203124581>
- Véliz, C. (2019). Three things digital ethics can learn from medical ethics. *Nature Electronics*, 2(8), 316–318. <https://doi.org/10.1038/s41928-019-0294-2>
- Williams, J. (2018). *Stand Out of Our Light: Freedom and Resistance in the Attention Economy*. Cambridge University Press.