Digital Arms for Digital Consumer Harms

Summary

Today, e-commerce is more popular than ever. It may also be more problematic than ever for consumers. Online traders have been found to use dark patterns – user interface design elements that may influence consumers into making choices that serve the business's bottom line. While in the EU we already have instruments that protect consumers from traders' exploitative conduct – the Unfair Commercial Practices Directive and Consumer Rights Directive, the proliferation of dark patterns in digital environments, including e-commerce websites, poses a threefold threat to the effectiveness of our current system of protection. First, many dark patterns operate on a behavioural dimension, exploiting our cognitive shortcomings, whereas our current system mostly acknowledges the informational dimension of consumer manipulation, and lends protection to rather rational average consumers in the form of information remedies. There thus seems to be a discrepancy between the kind of threats consumers encounter in digital environments and the sort of harms we regulate, as well as between the sort of consumers the law protects and real consumers. Second, we have designed this system to protect consumers both when transacting offline and online, i.e. our legal instruments are relatively technologically neutral. At the same time, the rates of compliance with these instruments in digital environments have historically been – and remain – worryingly low, and it cannot be ruled out that the technologically-neutral design of our current regulatory framework is not sufficiently clear on how to design compliant digital products such as user interfaces. Third, there is a mismatch between the scale at which dark patterns are deployed and the scale at which the enforcement of consumer protection instruments in digital markets by public authorities takes place. It might be time for enforcement authorities to step up their digital market monitoring game by deploying technical solutions (computational infringement detection methods) against socio-technical sources of consumer harms such as dark patterns. Against this background, this thesis explores **how we can regulate dark patterns effectively within the realm of EU consumer protection law, considering both legal and technical solutions.**

To answer this overarching question, this thesis first looks at whether dark patterns need to be regulated. This question is tightly linked to how dark patterns are problematised. **Chapter 2** frames dark patterns as a socio-technical artefact, i.e. one of the possible products of user experience optimisation, a process subject to technical, economic and organisational considerations. **Chapter 3** shows that dark patterns may also be a vehicle for the behavioural exploitation of users of various digital products. The behavioural exploitation of consumers as a means of market manipulation is a well-known problem, and the need to regulate dark patterns can therefore be assessed based on well-established grounds for intervention in consumer markets. **Chapter 4**

looks at dark patterns through two frames that may justify intervention: (behavioural) law and economics and autonomy theory. Both of these frames support intervention, but the former cautions us that the selection of dark patterns to be regulated ought to be driven by empirical evidence of welfare costs.

If we agree that (some) dark patterns need to be regulated, the question of how we may do that arises. Combining insights from (behavioural) law and economics, autonomy theory and the regulatory branch of law and technology literature, **Chapter 4** argues that in terms of regulatory design, the effective regulation of dark patterns may require recourse to incremental, risk-based regulation through (more) technology-specific standards. Effective regulation in socio-technical environments that are marked by a high speed of (potentially harmful) innovation also seems to call for the set-up of mechanisms that can ensure the adaptability of regulation to changing landscapes of consumer harms. Regulation in such settings faces an ongoing risk of regulatory disconnection, which may compromise its effectiveness. Regulatory disconnection can be normative – i.e. a mismatch between the values and goals underlying the legal system and the kinds of harms that occur, or descriptive – i.e. a mismatch between the image of technology we adopt in regulation and how it actually operates or is developed.

Chapter 5 and **Chapter 6** cast the spotlight on our current (substantive) system of protection to gauge its effectiveness – in light of its preoccupation with the informational dimension of (average) consumer manipulation and its relatively technologically-neutral set-up – at tackling dark patterns. While this technologically neutral set-up means that our current framework is flexible enough to apply to a wide range of dark patterns, barring cases of untruthful or entirely missing information, it lends itself to a wide range of interpretations when it comes to translating its requirements into the design of digital products such as consumer-facing user interfaces of e-commerce websites, which could lead to its weaponisation by well-resourced, ill-intended regulatees, and unintended breaches by less-resourced, well-intended market players. This points to a descriptive disconnection. Further, the instruments' preoccupation with the protection of average consumers through information remedies entails that they may miss the mark when it comes to protecting real, imperfectly rational consumers that may be behaviourally exploited by traders using dark patterns. This means that we may be dealing with a normative disconnection. We are also yet to adopt mechanisms that would ensure the long-term effectiveness of our regulatory scheme by enabling it to bounce back from regulatory disconnection.

Chapter 7 looks at how we may address these potential shortcomings in consumer protection in digital markets, and formulates some recommendations for the design of policy addressing traders' use of dark patterns. I suggest that rather than using the Commission's ongoing Digital Fairness Fitness Check of EU consumer law instruments as an opportunity to review the foundations of the current protective system, we could start small(er) by regulating some dark patterns that are well known to cause consumers

detriment in a technology-specific manner, through direct prescriptions or prohibitions of certain user interface design elements. We may also want to give our regulatory environment the necessary tools to adapt to changing socio-technical landscapes of consumer harms; these tools could take the form of regular evaluations, empowering the Commission to further specify (through implementing acts) or adjust legislative acts (through delegated acts), or recourse to New Approach co-regulation.

Whatever policy we adopt, it still needs to be enforced, however, and effective enforcement presupposes the timely detection of unlawful practices on digital markets. In **Chapter 8,** I explore whether and how enforcement authorities may use computational (web measurement) methods to monitor the deployment of unlawful dark patterns on e-commerce websites, and how regulatory design (the technology neutrality of substantive policy) may affect the technical feasibility of this exercise. In good news, the technological state of the art means that it is possible, in some cases, to automatically collect data about dark patterns from e-commerce websites at scale and to automatically map extracted textual and design elements to dark patterns and legal infringements. The use of these methods comes with some challenges and limitations, however, and some of these challenges may be related to the technologically neutral shape of policies. Technology-neutral regulation entails design variability, which complicates data collection efforts; it may lead to data inaccessibility and design volatility, which may result in the deprecation of data collection and analysis methods; it may also make it difficult to derive measurable infringement markers from policies, which presents a significant obstacle for the automation of data analysis. In other words, in digital environments, we may need technology specificity to both increase the effectiveness of legal solutions to the proliferation of dark patterns, as well as to enable the development of technical solutions that can support the detection of unlawful dark patterns at scale.

From an academic perspective, this thesis represents a contribution to the ongoing discussion in EU consumer law scholarship on how to regulate the use of dark patterns, and to the regulatory branch of law and technology literature. The findings of this thesis are also relevant for EU policymakers and national enforcement authorities.

Samenvatting

Titel: Digitale wapens voor digitale consumentenschade: Verkenning van juridische en technische oplossingen voor *dark patterns* in het consumentenrecht van de EU

Tegenwoordig is e-commerce populairder dan ooit. Voor consumenten kan het ook problematischer zijn dan ooit. Online handelaren blijken gebruik te maken van 'dark patterns': ontwerpelementen van de gebruikersinterface die consumenten ertoe kunnen aanzetten keuzes te maken die de bedrijfsresultaten dienen. Terwijl we in de EU al over instrumenten beschikken die consumenten beschermen tegen exploitatief gedrag van handelaars (de Richtlijn oneerlijke handelspraktijken en de Consumentenrechtenrichtlijn), vormt de proliferatie van *dark patterns* in digitale omgevingen, waaronder e-commercewebsites, een drievoudige bedreiging voor de doeltreffendheid van ons huidige beschermingssysteem. Ten eerste werken veel *dark patterns* op een gedragsmatige dimensie, waarbij ze onze cognitieve tekortkomingen uitbuiten, terwijl ons huidige systeem vooral de informatieve dimensie van consumentenmanipulatie erkent en bescherming biedt aan tamelijk rationele gemiddelde consumenten in de vorm van informatieoplossingen. Er lijkt dus een discrepantie te bestaan tussen het soort bedreigingen waarmee consumenten in digitale omgevingen te maken krijgen en het soort schade dat we reguleren, en tussen het soort consumenten dat door de wet wordt beschermd en echte consumenten. Ten tweede is dit systeem ontworpen om consumenten te beschermen bij zowel offline als online transacties; dat betekent dat onze juridische instrumenten relatief technologisch neutraal zijn. Tegelijkertijd zijn en blijven de nalevingspercentages van deze instrumenten in digitale omgevingen historisch gezien zorgwekkend laag, en kan niet worden uitgesloten dat het technologisch neutrale ontwerp van ons huidige regelgevingskader niet voldoende duidelijk is over hoe we digitale producten zoals gebruikersinterfaces kunnen ontwerpen die aan de regels voldoen. Ten derde is er sprake van een discrepantie tussen de schaal waarop *dark patterns* worden ingezet en de schaal waarop de handhaving van consumentenbeschermingsinstrumenten op digitale markten door overheden plaatsvindt. Het is misschien tijd voor toezichthouders om hun toezicht op de digitale markt op te voeren door technische oplossingen (computationele detectiemethoden voor het opsporen van overtredingen) in te zetten tegen sociaal-technische bronnen van consumentenschade, zoals *dark patterns*. Tegen deze achtergrond onderzoekt dit proefschrift **hoe *dark patterns* effectief kunnen worden gereguleerd binnen het kader van de Europese regels voor consumentenbescherming, waarbij zowel juridische als technische oplossingen in aanmerking worden genomen.**

Om deze overkoepelende vraag te beantwoorden, wordt in dit proefschrift eerst gekeken of *dark patterns* gereguleerd moeten worden. Deze vraag houdt nauw verband met de manier waarop *dark patterns* worden geproblematiseerd. **Hoofdstuk 2** beschrijft *dark patterns* als een sociaal-technisch artefact, dat wil zeggen een van de mogelijke resultaten van de optimalisatie van de gebruikerservaring, een proces dat onderhevig is aan technische, economische en organisatorische overwegingen. **Hoofdstuk 3** laat zien dat *dark patterns* ook een middel kunnen zijn voor gedragsmatige uitbuiting van gebruikers van verschillende digitale producten. De gedragsmatige uitbuiting van consumenten als middel voor marktmanipulatie is een bekend probleem, en de noodzaak voor het reguleren van *dark patterns* kan daarom worden beoordeeld op basis van bekende redenen voor interventie op de consumentenmarkten. **Hoofdstuk 4** kijkt naar *dark patterns* via twee kaders die interventie kunnen rechtvaardigen: (*behavioural*)rechts-economie en autonomietheorie. Beide kaders ondersteunen interventie, maar het eerste waarschuwt ons dat de selectie van *dark patterns* die moeten worden gereguleerd, moet worden gestuurd door empirisch bewijs van de welvaartskosten die erdoor worden veroorzaakt.

Als we het erover eens zijn dat (sommige) *dark patterns* gereguleerd moeten worden, rijst de vraag hoe we dat kunnen doen. Door inzichten uit de (*behavioural*) rechts-economie, de autonomietheorie en de regelgevende tak van de rechts- en technologieliteratuur te combineren, beargumenteert **Hoofdstuk 4** dat in termen van regelgevingsontwerp de effectieve regulering van *dark patterns* mogelijk een toevlucht vereist tot incrementele, op risico gebaseerde regulering door middel van (meer) technologiespecifieke standaarden. Effectieve regulering in socio-technische omgevingen die worden gekenmerkt door een hoge snelheid van (potentieel schadelijke) innovatie lijkt ook te pleiten voor het opzetten van mechanismen die het aanpassingsvermogen van de regulering aan veranderende landschappen van consumentenschade kunnen garanderen. Regulering in dergelijke situaties wordt geconfronteerd met een voortdurend risico van *regulatory disconnection*, wat de doeltreffendheid ervan in gevaar kan brengen. *Regulatory disconnection* kan normatief zijn – een discrepantie tussen de waarden en doelstellingen die ten grondslag liggen aan het rechtssysteem en de soorten schade die zich voordoen, of descriptief – een discrepantie tussen het beeld van technologie dat we bij de regelgeving aannemen en hoe deze feitelijk functioneert of wordt ontwikkeld.

Hoofdstuk 5 en **Hoofdstuk 6** focussen op ons huidige (materiële) beschermingssysteem om de effectiviteit ervan te meten – in het licht van de preoccupatie van dit system met de informatieve dimensie van (gemiddelde) consumentenmanipulatie en de relatief technologisch neutrale opzet – bij het aanpakken van *dark patterns*. Hoewel deze technologisch neutrale opzet betekent dat ons huidige raamwerk flexibel genoeg is om van toepassing te zijn op een breed scala aan *dark patterns*, behoudens gevallen van onwaarheid of geheel ontbrekende informatie, leent het zich voor een breed scala

aan interpretaties als het gaat om het vertalen van de eisen in het ontwerpen van digitale producten, zoals op de consument gerichte gebruikersinterfaces van e-commercewebsites, wat zou kunnen leiden tot bewapening ervan door goed uitgeruste, maar kwaadwillende marktactoren, en onbedoelde inbreuken door minder bemiddelde, goedbedoelende marktspelers. Dit wijst op een descriptieve *disconnection*. Bovendien brengt de preoccupatie van de Richtlijn oneerlijke handelspraktijken en de Consumentenrechtenrichtlijn met de bescherming van de gemiddelde consument door middel van informatiemiddelen met zich mee dat deze instrumenten de plank misslaan als het gaat om de bescherming van echte, imperfect rationele consumenten die gedragsmatig kunnen worden uitgebuit door handelaars die gebruik maken van *dark patterns*. Dit betekent dat we mogelijk te maken hebben met een normatieve *disconnection*. Daarnaast is de introductie van mechanismen nodig die de effectiviteit van het regelgevingssysteem op de lange termijn kunnen garanderen door het system in staat te stellen *regulatory disconnection* zelf te overkomen

Hoofdstuk 7 bekijkt hoe we deze potentiële tekortkomingen in de consumentenbescherming op digitale markten kunnen aanpakken, en formuleert enkele aanbevelingen voor het ontwerp van beleid dat het gebruik van *dark patterns* door handelaren aanpakt. Ik stel voor dat, de lopende *Digital Fairness Fitness Check* van de instrumenten van het EU-consumentenrecht van de Europese Commissie niet wordt gebruikt als een kans om de fundamenten van het huidige beschermingssysteem te herzien, maar dat in plaats daarvan klein(er) wordt begonnen door een aantal *dark patterns* te reguleren waarvan bekend is dat ze consumenten benadelen op een technologiespecifieke manier, door directe voorschriften of verboden op bepaalde ontwerpelementen van de gebruikersinterface. Misschien is het verder aan te raden om ons reguleringssysteem ook de nodige instrumenten geven om zich aan te passen aan de veranderende sociaal-technische landschappen van consumentenschade; deze instrumenten zouden de vorm kunnen aannemen van regelmatige evaluaties, waardoor de Commissie de bevoegdheid zou krijgen om rechtshandelingen verder te specificeren (door middel van uitvoeringshandelingen) of aan te passen (door middel van gedelegeerde handelingen), of een beroep te doen op coregulering volgens de Nieuwe Aanpak.

Welk beleid ook wordt gevoerd, het moet nog steeds worden gehandhaafd, en effectieve handhaving veronderstelt de tijdige detectie van onwettige praktijken op digitale markten. In **Hoofdstuk 8** onderzoek ik of en hoe toezichthouders computationele (*web measurement*) methoden kunnen gebruiken om de inzet van onwettige *dark patterns* op e-commercewebsites te monitoren, en hoe de design van regelgeving (de technologieneutraliteit van materieel beleid) de technische haalbaarheid van deze exercitie kan beïnvloeden. Het goede nieuws is dat de technologische stand van zaken betekent dat het in sommige gevallen mogelijk is om automatisch op grote schaal gegevens over *dark patterns* van e-commercewebsites te verzamelen en de geëxtraheerde tekst- en ontwerpelementen automatisch toe te wijzen aan *dark patterns* en wetschendingen. Het

gebruik van deze methoden brengt echter enkele uitdagingen en beperkingen met zich mee, en sommige van deze uitdagingen kunnen verband houden met de technologisch neutrale vorm van beleid. Technologieneutrale regelgeving brengt variabiliteit in de design met zich mee, wat de inspanningen voor het verzamelen van gegevens bemoeilijkt; het kan leiden tot ontoegankelijkheid van gegevens en volatiliteit in de design, wat kan resulteren in de veroudering van methoden voor het verzamelen en analyseren van gegevens; het kan het ook moeilijk maken om meetbare inbreukmarkers uit beleid af te leiden, wat een aanzienlijk obstakel vormt voor de automatisering van data-analyse. Met andere woorden: in digitale omgevingen hebben we mogelijk technologiespecificiteit nodig om zowel de effectiviteit van juridische oplossingen voor de proliferatie van *dark patterns* te vergroten, als om de ontwikkeling van technische oplossingen mogelijk te maken die de detectie van onwettige *dark patterns* op schaal kunnen ondersteunen.

Vanuit academisch perspectief vertegenwoordigt dit proefschrift aldus een bijdrage aan de voortdurende discussie in de Europese consumentenrechtwetenschap over hoe het gebruik van *dark patterns* kan worden gereguleerd, en aan de regelgevende tak van de recht- en technologieliteratuur. Daarnaast zijn de bevindingen van dit proefschrift ook relevant voor EU-beleidsmakers en nationale toezichthouders.