What is conceptual disruption?

Marchiori, S.; Scharp, Kevin

DOI
10.1007/s10676-024-09749-7

Publication date
2024

Document Version
Final published version

Published in
Ethics and Information Technology

Citation (APA)
Marchiori, S., & Scharp, K. (2024). What is conceptual disruption? *Ethics and Information Technology*, *26*, Article 18. https://doi.org/10.1007/s10676-024-09749-7

Ethics and Information Technology (2024) 26:18
https://doi.org/10.1007/s10676-024-09749-7

ORIGINAL PAPER

What is conceptual disruption?

Samuela Marchiori[1] · **Kevin Scharp**[2,3]

Accepted: 29 January 2024

Abstract
Recent work on philosophy of technology emphasises the ways in which technology can disrupt our concepts and conceptual schemes. We analyse and challenge existing accounts of conceptual disruption, criticising views according to which conceptual disruption can be understood in terms of uncertainty for conceptual application, as well as views assuming all instances of conceptual disruption occur at the same level. We proceed to provide our own account of conceptual disruption as an interruption in the normal functioning of concepts and conceptual schemes. Moreover, we offer a multilevel taxonomy thereof, where we distinguish between instances of conceptual disruptions occurring at different levels (conceptual scheme, conceptual clusters, and individual concepts), taking on different forms (conceptual gaps and conceptual conflicts), and leading to different degrees of severity (extending from mild to severe). We also provide detailed accounts through historical examples of how conceptual gaps and conceptual conflicts can occur at different times in the very same process of conceptual disruption. Finally, we make the case that different kinds of conceptual engineering can provide meaningful ways to assess and overcome distinct types of conceptual disruption.

Keywords Conceptual disruption · Conceptual gap · Conceptual conflict · Conceptual change · Conceptual overlap · Socially disruptive technologies

Introduction

Recent work on philosophy of technology emphasises not only the ways in which technology can disrupt our social structures, e.g., by making certain occupations obsolete, but also the ways in which it can disrupt our concepts (Hopster, 2021). The newly emerging literature on the matter considers *conceptual disruptions* as challenges or interruptions to the normal course of conceptual activity (Hopster, 2021; Hopster & Löhr, 2023; Hopster et al., 2023; Löhr, 2022, 2023a, 2023b). The same literature recognises three kinds of conceptual disruption: *gaps* (i.e., no concepts apply to some new phenomenon), *overlaps* (i.e., multiple incompatible concepts apply to some item), and *misalignments* (i.e., concept activity is not aligned with norms and values). Here we focus our attention on gaps and overlaps, while leaving the investigation of conceptual misalignments for future work.

We present a new tripartite taxonomy for conceptual disruption, which divides this phenomenon across types of disruption, levels of disruptions, and degrees of severity of the disruption. First, we introduce a new type of conceptual disruption, namely, a conceptual conflict. We suggest that the category of overlaps is better understood as conceptual conflicts, and we propose several benefits of making this change. Second, we emphasise the diversity in the category of conceptual disruptions by tracing out three distinct levels at which they occur: individual concepts, clusters of concepts, and conceptual schemes. Third, we distinguish among mild, moderate, and severe disruptions, based on the varying degrees of urgency for bridging instances of conceptual disruption.

We ground the taxonomy in a novel and substantive account of conceptual disruption and its types by appealing to the idea that concepts have constitutive principles that partly determine their identity. These accounts illuminate the

✉ Samuela Marchiori
s.marchiori@tudelft.nl

Kevin Scharp
kscharp@illinois.edu

[1] Department of Values, Technology and Innovation, Delft University of Technology, Delft, The Netherlands

[2] Department of Philosophy, University of Illinois Urbana-Champaign, Urbana, IL, USA

[3] Faculty of Behavioural, Management and Social Sciences, University of Twente, Enschede, The Netherlands

Published online: 07 March 2024

nature of conceptual disruption in general and the specific kinds of conceptual disruptions, but also the complex relationships between conceptual gaps, conceptual conflicts, and the conceptual changes that occur when technologies lead to changes in our concepts. Although we focus on technology-induced instances of conceptual disruption, the phenomena we consider can occur in a wider class of domains beyond the technological one.

The paper is structured as follows. "Background on conceptual disruption" introduces and engages with the literature on conceptual disruption and sets the stage for the taxonomy to follow. In "Conceptual disruptions and constitutive principles", we present both the view that concepts have constitutive principles and the idea that concepts have certain functions. Such insights are applied in "Conceptual disruption: a framework", which proposes a tripartite framework for conceptual disruption. Here, we distinguish among conceptual disruptions occurring at three levels (individual concepts, clusters of concepts, conceptual schemes), taking on two forms (conceptual gaps, conceptual conflicts), and leading to three different levels of severity (mild, moderate, severe). Furthermore, we avail ourselves of two case studies to illustrate how conceptual gaps and conceptual conflicts can occur at different times in the very same process of conceptual disruption, and how they lead to different kinds of conceptual changes. "The role of conceptual engineering" focuses on the relationship between conceptual disruption and conceptual engineering, which is the process of evaluating our concepts and improving our concepts and conceptual schemes. "Conclusion" contains our conclusions.

Background on conceptual disruption

The philosophical literature on conceptual disruption is currently an emerging research niche, as the phenomenon only recently gained attention from philosophers of technology, in particular scholars interested in the ethics of socially disruptive technologies such as artificial intelligence, blockchain, and gene editing (Hopster, 2021; Hopster & Löhr, 2023; Hopster et al., 2023; Löhr & Michel, 2023; Löhr, 2022, 2023a, 2023b). Indeed, new technologies can change the way our concepts function or the features of our conceptual schemes, leading to conceptual disruptions.

For the purpose of this paper, we will adopt Machery's (2009) definition of concepts, whereby "a *concept* of x is a body of [information][1] about x that is stored in long-term memory and that is used by default in the processes underlying most, if not all, higher cognitive competences when these processes result in judgments about x" (Machery, 2009, p. 12, our italics).[2] We do not distinguish between concepts and linguistic meanings for the purposes of this paper.

In recent work representative of the nascent conceptual disruption literature as a whole, conceptual disruptions are defined as "a challenge to the meaning of concept that prompts a possible future revision of it" (Hopster et al., 2023, p. 143).[3] In addition, conceptual disruptions are taxonomised in the following way:

> We argued that conceptual disruptions can be interpreted in three ways, and that technology typically plays a prominent role in each of them (Hopster and Löhr, 2023). First, we may be faced with a 'conceptual gap'. That is, we lack the concepts needed to describe a novel technological artifact, or to normatively evaluate the new impacts and affordances to which it gives rise. Second, we may be faced with a 'conceptual overlap'. That is, more than one of our existing concepts may be appropriate to describe and evaluate a novel technology, but there is uncertainty as to which concept is most suitable. Third, there may be cases of 'conceptual misalignment'. In such cases, existing concepts do seem applicable to conceptualize a new technology and its impacts and affordances. However, this apparent good fit actually masks an underlying value misalignment: the concept and its use do not express the values that a community of concept-users, upon ethical reflection, would like it to express. (Hopster et al., 2023, p. 144)

In what follows, we focus on only the first two kinds of disruptions, gaps and overlaps; we leave misalignments for future work. The ways in which each of these three terms has been defined thus far do not fully capture some features that we deem significant.

First, conceptual disruptions are not best thought of as consisting in the challenging of the meaning of a concept. For example, when there is a conceptual gap, which is a kind of conceptual disruption, there is some phenomenon to which our existing concepts cannot properly be applied. It might be that the phenomenon in question does not qualify as belonging to the extension of any of our existing concepts. In this case, no concept has its meaning challenged. Rather, the conceptual scheme as a whole is deemed inadequate for the purposes at hand. Thus, the above definition of

[1] In the original formulation, "body of knowledge" (Machery, 2009, p. 12).

[2] It should be noted that Carey's (2009) work on concepts aligns with this definition and she provides a vast amount of detail for each of these aspects and embeds them within a much larger and empirically-supported framework for thinking about the relationship between concepts and other kinds of mental phenomena.

[3] It should be noted that, while we choose to focus on this contribution, this is not the first publication on the matter, as evidenced by the papers cited above.

"conceptual disruption" does not need to adequately capture the phenomenon under investigation.

Löhr's work on conceptual disruption emphasises the *conceptual uncertainty* associated with disruptions to our concepts. For example, Löhr writes:

> The term *conceptual disruption* in this paper means: Any intentional or unintentional challenge or interruption of the ways in which the individual or group has intuitively classified individuals, properties, actions, situations, or events, leading to *classificatory uncertainty*, i.e., uncertainty about the application conditions of a word or concept," (Löhr, 2022, p. 838, our italics).

Admittedly, Löhr is not proposing a definition that holds in general, but merely making a stipulation to capture the phenomenon for the purpose of the contribution under consideration. Nevertheless, since this is one of the few papers on the subject, we are interested in understanding whether this may hold, were this to be treated as a more general proposal for a definition of conceptual disruption. However, it should be noted that the following criticism might not apply to Löhr.

We have two worries about this stipulation. First, it applies only to classification, which is a function performed by concepts expressed by *predicates* in language. However, conceptual disruption is more widespread. For example, the development of quantum mechanics and its counterintuitive consequences caused some to doubt the distribution law in classical logic, i.e., A or (B and C) iff (A and B) or (A and C). This doubt about distribution led to the development of quantum logics that violate the distribution law (Putnam, 1965). This is clearly a conceptual disruption, but it pertains to logical connectives, not predicates used for classification.[4]

The second worry about the definition under consideration is that, although we agree that uncertainty about whether certain concepts properly apply to a case does often accompany conceptual disruptions, it does not seem to occur in every case. For example, the concept EXOPLANET and the term "exoplanet" go back to at least 1992 when astronomers hypothesized the existence of objects that are similar to planets but orbit other stars. Planets have been defined as orbiting the sun, so these new objects were clearly not planets and no other astronomical concept seemed to apply to them. Thus, a new concept and new term were invented to classify this phenomenon. In this example, it does not seem like there was any uncertainty about whether the objects in question were planets. Everyone agreed that they were not planets. Nor was there uncertainty about whether these objects would be moons or some other kind of familiar celestial object. Instead, it was clear that a new category was needed to describe these objects properly (Burke, 1992). Therefore, while instances of conceptual disruption may oftentimes involve a certain degree of classificatory uncertainty, such uncertainty is not a necessary feature of conceptual disruption as a whole.

Moreover, it seems clear that there are cases of conceptual uncertainty that are not conceptual disruptions. For example, in 2011, some physicists working on the Oscillation Project with Emulsion-tRacking Apparatus (OPERA) project measured the velocity of some neutrinos as being greater than the speed of light. This caused considerable uncertainty about whether the concept of superluminal motion applied to these neutrinos. It turned out that a fibre optic cable was not attached properly to a piece of equipment which resulted in the mistaken measurement (Cartlidge, 2012). Therefore, while instances of conceptual disruption may manifest as instances of classificatory uncertainty, conceptual uncertainty is neither necessary nor sufficient for conceptual disruption. Instead, we propose that an interruption in the normal functioning of a concept, cluster of concepts, or conceptual scheme, is a necessary condition for disruption.

Finally, we disagree with how one kind of conceptual disruption, conceptual overlaps, have been characterised. Hopster et al. (2023) offer the following definition: "[A conceptual overlap emerges if] more than one of our existing concepts may be appropriate to describe and evaluate a novel technology, but there is uncertainty as to which concept is most suitable" (p. 144).

One worry is that the term "overlap" does not fully capture the problem with this kind of conceptual disruption in a perspicuous way. Indeed, there is an assumption that conceptual overlaps involve the application of *incompatible* concepts (e.g., natural and artificial, or alive and dead) (Hopster & Löhr, 2023). However, referring to "overlaps" may lead to misunderstandings. The fact that two or more concepts overlap is not a problem by itself. For example, the concepts BROWN and DOG overlap on brown dogs in the sense that brown dogs are in the extension of the predicate "brown" and in the extension of the predicate "dog". Of course, this situation does *not* involve a conceptual disruption. Thus, the term "overlap" may not be the best choice to capture such a kind of conceptual disruption.

Moreover, as we will see, this kind of conceptual disruption can occur with a single concept, so the term "overlap" is apt to be misleading. We prefer the term "conflict" instead to capture the sense of incompatibility at

[4] An anonymous referee pointed out that logical connectives can be used for classification of sentences, which is true. However, they also have non-classificatory uses (e.g., uttering a conjunction), and that is all we need for the counterexample.

play in these examples.[5] In addition, the kind of disruption in question involves some sort of conflict among concepts or between concepts and facts in the world. Thus, our category of conceptual conflicts incorporates what theorists have called conceptual overlaps, but is somewhat more expansive. We expand on this point in the next section.

Conceptual disruptions and constitutive principles

In this section, we outline a theory of conceptual disruption in general and what distinguishes particular kinds of conceptual disruptions like conflicts and gaps. The key element in the theory we offer is that concepts are determined by constitutive principles which help guide the application of the concept and provide identity conditions for concepts.

A constitutive principle for a concept is part of the content of that concept—they are "built in" to concepts. Constitutive principles are descendants of analytic principles. However, constitutive principles are designed to be different from analytic principles in at least two ways. First, constitutive principles can be false. This aspect marks an appreciation of the fact that our concepts and the meanings of our words are not pristine. They are historically situated and potentially messy items that have worked well enough in past situations for our ancestors to have passed them down to us. Concepts with false constitutive principles might work well in most situations, like our concepts of absolute space and absolute time. We know these concepts incorporate false claims about the world (e.g., simultaneity is observer-independent), but they do not cause problems in everyday situations, so it is fine to use these concepts in those contexts. Second, constitutive principles are not mandatory beliefs. That is, one need not believe that all bachelors are unmarried in order for one's word 'bachelor' to mean BACHELOR. Instead, the relationship between the possessors of a concept and that concept's constitutive principles is more complex (Eklund, 2002).

The theory of constitutive principles outlined here is based on previous work from one of the authors (Scharp, 2013). When one has a conversation and one's interlocutor denies a constitutive principle, that is an "interpretive red flag"—this event ought to give one pause about going on interpreting the interlocutor in the usual way. A philosophically substantive way of explaining this metaphor is to say:

> X is a *constitutive principle* for concept C iff when an interlocutor denies X, one has a pro tanto reason to think that one does not mean the same thing as one's interlocutor by the word that expresses C.

Because the sentence "dogs are animals" is constitutive for the concept DOGS, if a person speaking to you sincerely asserted that dogs are not animals, that would give you a defeasible reason to believe that this person does not mean what you mean by DOG. The reason is defeasible, so it can be overcome. That is, one might discover that this person does mean what you mean by DOG but simply has very strange or unusual beliefs about dogs. Moreover, one who possesses concept C is *entitled* to the constitutive principles for C in the sense that one would be justified in believing them unless one has strong evidence to the contrary. So anyone who possesses the concept of dog is entitled to believe that dogs are animals, even though (presumably) some people who possess this concept do not have this belief.

One might reasonably wonder how a constitutive principle can be false it if it implicitly defines the term in question. This aspect can be confusing, especially because of the dominance of analyticity in thinking about meaning and concepts. A constitutive principle is just any old proposition with the extra status that we who possess and use the concept in question are going to presuppose that the principle in question is true and we are going to expect others to presuppose this as well. However, what we presuppose to be true can turn out to be false. Moreover, it might seem like stipulative definitions could never be false, because they give the word in question whatever meaning makes them true. However, that cannot be right since some stipulative definitions are impossible to make true (e.g., !tonk"#like definitions). Instead, a stipulative definition gives the term in question whatever meaning makes the definition constitutive, which is about how one thinks and communicates.

With this understanding of constitutive principles, we can provide substantive accounts of the kinds of conceptual disruptions and how they relate to conceptual change and conceptual engineering in particular. A disruption is usually understood as an interruption in the normal course of a process or activity. As such, a *conceptual disruption* is an interruption in the normal functioning of a concept, cluster of concepts,[6] or conceptual scheme (in this sense, also Löhr, 2023a).

[5] It should be noted that, within the growing body of literature on conceptual disruption, different authors may attribute different meanings to the same terms. Therefore, the focus of attention should be placed not on the terms themselves, but on the way in which such terms are conceptualised.

[6] What we mean by conceptual cluster is a set of concepts that are interrelated. For example, ALIVE and DEAD are part of a cluster of concepts because they are part of a network of interlinked concepts. This is relevant as changes to the former may force changes in the latter due to the concepts' interconnectedness. Conversely, ALIVE and NEUTRINO are not part of the same cluster because they have very little to do with one another. Conceptual clusters can vary in size, but are always smaller than a conceptual scheme.

Table 1 Taxonomy of conceptual disruption

	Level of disruption		Type of disruption	Severity of disruption
Conceptual disruption	Level 1	Conceptual scheme	Conceptual gap	Mild disruption
	Level 2	Conceptual cluster	Conceptual gap	Moderate disruption
			Conceptual conflict	Severe disruption
	Level 3	Individual concept	Conceptual gap	Moderate disruption
			Conceptual conflict	Severe disruption

In this definition, we have appealed to the normal functioning of a concept or concepts. By this, we just mean the pattern of using the concept for certain purposes in certain situations. It is common in the philosophical literature on conceptual engineering (i.e., the philosophical study of how to assess and improve our conceptual resources) to appeal to the functions of concepts in order to explain why a new concept is a good replacement for an existing concept (for a survey, see Riggs, 2021). However, this appeal has generated considerable controversy. On the one hand, this appeal to conceptual functions is popular, but on the other hand there is also strong opposition to it (Cappelen, 2018). The debate has generated a host of new substantive accounts of what conceptual functions are and some philosophers have advocated minimalist views on the matter. For example, Jorem argues that advocates of conceptual engineering ought to say no more than that a function of a concept is simply something the concept is used for (Jorem, 2022). We do not take a stand on this issue here, but appealing to conceptual functions allows us to draw important distinctions between conceptual disruption and conceptual change.

Conceptual disruption can be identified as the precursor to a conceptual change, whereby the interruption brought about by the conceptual disruption may then be followed by conceptual change.[7] In this sense, while conceptual disruption itself is a kind of change, it is not a kind of *conceptual* change. A conceptual change is the addition or elimination of a concept or a change in one or more constitutive principles for one or more concepts. A conceptual disruption is a change in the normal functioning of a concept, cluster of concepts, or conceptual scheme. Although conceptual changes are often prompted by conceptual disruptions, the two phenomena are distinct (Löhr, 2023a).

Conceptual disruption: a framework

It should be noted that conceptual disruption is not a neatly ordered group of phenomena, rather a fairly heterogeneous set of phenomena that ought to be explained to some extent independently from each other.

It is important to emphasise that the phenomenon of conceptual disruption should better be understood as occurring at different levels of granularity. In some instances, conceptual disruption primarily affects one or more concepts. In others, it primarily affects the conceptual scheme as a whole. Glossing over this different way of framing instances of conceptual disruption can make it difficult to understand the phenomenon itself, and may risk attributing features to conceptual disruption in general that apply only to certain kinds of them.

Indeed, conceptual disruption is a multi-layered phenomenon occurring on three main levels (conceptual schemes, conceptual clusters, individual concepts), which manifests in a number of distinct forms (conceptual gaps and conceptual conflicts), and presents varying degrees of severity (extending from mild to severe) (Table 1).

Levels of conceptual disruption

The first level of conceptual disruption affects the conceptual scheme. At this level, the conceptual disruption is a change in a property of a conceptual scheme, such that the conceptual scheme that was complete before the instance of conceptual disruption occurred is no longer exhaustive. That is to say, a conceptual scheme that previously applied to almost any experienced phenomenon and could adequately classify them, is no longer able to do so.[8]

For example, consider physicists inventing a new material,[9] biologists discovering a new species, engineers developing a new technology, or astronomers discovering

[7] In this sense, conceptual disruption and conceptual change could be considered as two components of a (micro) conceptual revolution, as described in Thagard (1993). However, the analogy is not perfect, as conceptual disruption is not a rare occurrence, nor does it require the equivalent of a paradigm shift at the macro level. Rather, it occurs regularly and, as illustrated in "The role of conceptual engineering", it can be addressed and bridged within the confines of "normal science" i.e., conceptual engineering.

[8] It should be noted that, as an anonymous referee pointed out, conceptual schemes are hardly ever comprehensive. However, non-exhaustive schemes may appear to be exhaustive until a conceptual disruption uncovers their lack of completeness.

[9] For instance, see 2DPA-1, the material created in 2022 by MIT engineers (Zeng et al., 2022).

a new celestial object. In any of these scenarios, neither our concepts nor our conceptual scheme change in terms of something being added or subtracted. What does change is the property of the scheme as being comprehensive. This is a change in the function of the conceptual scheme, not a change in any constitutive principles for the concepts in question.

The second level of conceptual disruption affects conceptual clusters. In the second case, the instance of conceptual disruption affects multiple concepts at once, which may trigger additional disruptions in related concepts. An example is the disruption of the concepts LIFE and DEATH following the introduction of the mechanical ventilator (Baker, 2013; Nickel, 2020; Nickel et al., 2022).

The third level of conceptual disruption affects individual concepts. In this case, the instance of conceptual disruption affects one concept, which then oftentimes triggers a disruption in related concepts. An example is the disruption of the concept of moral agency following the introduction of artificial intelligence (e.g., Brożek & Janik, 2019; Sullins, 2009). However, the disruption is not limited to one concept, but leads to more concepts linked to the first one also being disrupted (e.g., moral responsibility, as in Babushkina, 2020; Santoni de Sio & Mecacci, 2021; Sebastián, 2021), somewhat like a domino effect. In certain instances, the latter set of concepts would not necessarily have been disrupted if the former concept had not been disrupted to begin with. This echoes the notion of "inferential risk" in Löhr and Michel (2023).

In the second and third level, we can further distinguish between first-order disruptions and second-order disruptions. Indeed, our concepts are so interconnected with one another that conceptual disruption affecting one or more concepts is likely to ramify and propagate through the conceptual scheme. As a result, instances of conceptual disruption can exert direct effects on one or more concepts, and indirect effects on nearby concepts. For example, by blurring the boundaries between the human and the non-human, social robots can be said to disrupt the concept PERSON, which can prompt the disruption of AGENCY, which in turn may prompt the disruption of RESPONSIBILITY, and so on (van de Poel et al, 2023). We adopt the expressions *first-order disruption* and *second-order disruption* to account for the difference between concepts being directly and indirectly affected by the conceptual disruption, respectively.

In addition to characterising conceptual disruption, we focus on two specific types: gaps (in which the disruption affects the feature of exhaustiveness of concepts and conceptual schemes) and conflicts (in which the disruption manifests itself as an incompatibility among different concepts or constitutive principles). Conceptual gaps can take place at the level of the conceptual scheme, at the level of the conceptual cluster, and at the level of the individual concept. Conceptual conflicts can occur at the level of individual concepts and conceptual clusters. Let us illustrate each type of conceptual disruption in more detail.

Types of conceptual disruption

Conceptual gaps

A *conceptual gap* occurs if and only if there is some phenomenon that ought to be classified at a certain level of precision but no relevant concepts properly apply to the phenomenon in question. The phrase "at a certain level of precision" is crucial in this definition since many phenomena fit into categories at multiple levels of description.

For example, in the case of exoplanets described above, scientists were not completely without ways of classifying exoplanets before this conceptual innovation. Everyone would agree that something like a planet that is orbiting a star other than the sun would be a physical object, and it would be an object roughly in hydrostatic equilibrium from its own gravity, and it would be an object in orbit, and so on. Hence, the concept of physical object and the others listed would properly apply to this phenomenon. Thus, it is not the case that we have *no way whatsoever* of classifying something that is a conceptual gap. Rather, it is that we do not have any way of classifying it *at the relevant level of precision.* Moreover, the "ought" in the definition of "conceptual gap" pertains to the relevant projects, goals, and aims of the people involved. In the case of exoplanets, there was a conceptual gap because scientists wanted a way to classify these celestial objects that would differentiate them from other kinds of celestial objects like asteroids, comets, and moons.

With regard to the first level of conceptual disruption, conceptual gaps highlight a feature of the conceptual scheme, such that the scheme was complete before the instance of disruption, and ceases to be so following the conceptual disruption. A conceptual scheme exhibits the feature of nonexhaustiveness whenever a new socio-technical phenomenon emerges, which none of the concepts within it are able to express. This type of disruption prompts the introduction of a new concept to fill such a gap and restore the exhaustiveness of the conceptual scheme. For example, the development of artificial intelligence technologies disrupted our conceptual scheme and prompted the subsequent introduction of the concept ARTIFICIAL INTELLIGENCE, which effectively appeased such a disruption.

As for the second level of conceptual disruption, nonexhaustiveness affects conceptual clusters in the sense that, while one or more existing concepts may allow to partially express a phenomenon, such concepts either do not always allow to do so to a satisfactory degree, or can convey the phenomenon satisfactorily, but highlight the need for a more

straightforward way to do so, e.g., through the introduction of a new concept.

An example of non-exhaustiveness as the need to satisfactorily convey a phenomenon is the case of PRIVACY. Despite PRIVACY being a somewhat recent concept (Warren & Brandeis, 1890), especially when compared to some of its neighbouring concepts (such as AUTONOMY, DIGNITY, FREEDOM), one could argue that a series of phenomena largely overlapping with our understanding of the phenomenon currently identified with PRIVACY occurred well before the introduction of the related concept (DeCew, 2018). Moreover, one could argue that privacy can be expressed to a considerable extent as a combination of different concepts. An example would be expressing PRIVACY as the intersection of "moral autonomy, human dignity, identity, personhood, liberty, anonymity and confidentiality" (van de Poel, 2020, p. 55). We argue that what prompted the introduction of PRIVACY was an instance of conceptual disruption in terms of non-exhaustiveness. That is, despite the combination of other concepts being somewhat able to express the phenomenon that PRIVACY would later convey, the advent of photography and its uses in sensationalist journalism highlighted that something was still missing: a new concept that would convey all the nuance of that particular phenomenon that needed to be expressed (Warren & Brandeis, 1890). In this scenario, the instance of conceptual disruption manifested itself as uncertainty in relation to both the nature of the conceptual gap and how such a gap should be bridged.

An example of non-exhaustiveness understood as the need for more straightforward ways to convey a phenomenon is the case of INFORMATION PRIVACY.[10] One could argue that INFORMATION PRIVACY can be adequately expressed as a combination of different concepts, e.g., PRIVACY, INFORMATION, DATA, INFORMATION TECHNOLOGY. Therefore, what prompted the introduction of the concept INFORMATION PRIVACY was an instance of conceptual gap such that, while the phenomenon under consideration could have been conveyed by only relying on existing concepts, its salience justified the development of a new concept to facilitate its expression.

What makes instances of conceptual disruption as non-exhaustiveness on the conceptual scheme level different from those occurring on the conceptual cluster level is ultimately a matter of degree. In the former case, the gap is such that it is extremely hard, if not impossible, to meaningfully express the new concept using the existing conceptual repertoire. In the latter case, uncertainty emerges over the presence, extension, and severity of the conceptual gap, as well as over the application conditions of the existing concepts.

As for the third level of conceptual disruption, conceptual gaps affect individual concepts whenever they do not seem able to adequately convey the phenomenon that they should express. While this does not always occur as a result of technological advancements, technology can play a role in undermining existing concepts, e.g., by uncovering potential issues with relation to assumptions or constitutive principles related to such concepts. In this case, the non-exhaustiveness manifests itself as a defect in a concept that makes it inadequate to express certain nuances that it should be able to encompass.

Conceptual conflicts

A *conceptual conflict* occurs if and only if there is some known fact that is inconsistent with some constitutive principles for some concepts.[11] For example, in the late 1990s and early 2000s, scientists using new astronomical technologies were able to identify great numbers of objects in the Kuiper belt that were very similar to Pluto. These objects seemed to satisfy the definition of PLANET that was used at the time, which goes back to antiquity.

Nevertheless, astronomers were reluctant to call these objects planets for at least two reasons. First, planets are supposed to be special in that there are very few of them. Second, there seemed to be a difference in kind between Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, and Neptune on the one and Pluto and the other Kuiper belt objects on the other hand. Lumping all these objects together under a single category obscured an important natural difference among them.

To remedy the situation, professional astronomers voted to redefine the word "planet" so that all planets have cleared the neighbourhood around their orbit of debris (IAU, 2006). This redefined concept of planet applied to all the familiar objects except Pluto. Hence, on the new definition, Pluto is not a planet. Moreover, the astronomers introduced a new concept and word—dwarf planet—that

[10] For the purpose of this paper, following Chalmers (2020), we consider "information privacy" to be a separate (while deeply interconnected) concept instead of a conception of "privacy".

[11] A clarification is in order. Conceptual conflicts can occur in three main ways: (a) as inconsistencies among constitutive principles; (b) as inconsistencies between constitutive principles and facts about the world; (c) as inconsistencies between constitutive principles and values or social norms. Here, we address the first two types, and leave the third for future work on conceptual misalignment. In this paper, we focus specifically on the second type of conflict, and characterise conflicts as inconsistencies between constitutive principles and facts. This is because it is extremely difficult to generate a conflict among constitutive principles alone, as there seems to always be some wiggle room. The liar paradox can exemplify this challenge: in order to get the contradiction, one first needs to have the empirical claim that the liar's sentence exists.

applies to Pluto and the other Kuiper belt objects in question. In this case, the conceptual conflict was an inconsistency between the constitutive principles for the concept of planet (i.e., there are few planets) and the empirical discovery that there are many objects orbiting the sun in roughly hydrostatic equilibrium from their own gravity (i.e., there are very many planets). There cannot be both only a few planets and not only a few planets. This is the heart of the conceptual conflict. The conflict in this case was resolved by adding a new constitutive principle in a redefinition of PLANET so that planets have cleared the debris from the neighbourhoods of their orbits.

Conceptual conflicts take place at the second and third level of conceptual disruption, affecting conceptual clusters and individual concepts. When they affect individual concepts, they manifest themselves as instances of incompatibility among a concept's constitutive principles. An example of this is the disruption of the concept TRUTH (Scharp, 2013). When they affect conceptual clusters, they manifest themselves as instances in which two or more concepts are only partially able to convey what we want to express, but none of them is able to do so to a satisfactory degree, or to a distinctly better degree compared to the other(s), which leads to a phenomenon being expressed through a combination of mutually incompatible concepts.

For example, in the case of a person whose cognitive functions are irrevocably lost, but who is able to breathe through mechanical ventilation, we might find it difficult or even impossible to determine whether either the concept ALIVE or DEAD can adequately convey what we want to express about the condition of the person. Not only does neither concept taken by itself seem good enough to express such a phenomenon, but the combination of the two concepts does not seem to meaningfully help solve the issue. That is to say, describing the condition of the person in the example as them being alive, or as them being dead, does not seem to fully convey the salient features of the phenomenon at hand. At the same time, describing the condition of the person as them being both alive and dead at the same time does not seem particularly helpful either, given that the state of being alive is incompatible with the state of being dead.

Whenever instances of conceptual disruption affect conceptual clusters, we are confronted with uncertainty as to how to determine the conditions according to which existing concepts can or should be deemed suitable to fill the gap of a seemingly missing concept, and apply such conditions to the case at hand. In such scenarios, it may be challenging to discern instances of conceptual gaps from instances of conceptual conflicts.

For example, in the case of the mechanical ventilator, one may argue that the conceptual disruption at hand is an instance of non-exhaustiveness, to the extent that neither the concept ALIVE nor the concept DEAD can adequately express what we want to convey about the status of a person who is able to breathe through a ventilator, but has no cognitive functions and cannot regain them. However, one may also argue that this is an instance of conceptual conflict, as both concepts are partially able to convey what we want to convey, and the difficulty lies not so much in determining *whether* one of those concepts or the combination of those concepts is adequate enough to describe the phenomenon at hand, but rather in us not being able to determine *which* of the two concepts, equally suitable for conveying what the seemingly missing concept should convey, should be preferred.

We argue that whether an instance of conceptual disruption should be described in terms of conceptual gap or conceptual conflict ultimately depends on the constitutive principles of the concepts involved. That is, understanding whether (the constitutive principles of) two (or more) concepts stand in an A versus B or A versus ¬A relation allows to qualify the conceptual disruption at hand as an instance of conceptual gap or conceptual conflict, respectively (in similar terms, Sommaggio & Marchiori, 2018).

Let us illustrate this with an example. In both the cases of PRIVACY and BRAIN DEATH, the instance of conceptual disruption emerged in such a way that a phenomenon could not be adequately addressed by any existing individual concept, and could only partially be addressed by a combination of existing concepts. However, the way in which such concepts relate to each other and come together to express the seemingly missing concept (PRIVACY and BRAIN DEATH, respectively) varies significantly.

In the case of privacy, the individual concepts considered in the combination stand against each other in an A versus B type of relation. That is, neither the concepts as a whole nor their constitutive principles are incompatible with each other, thus identifying a conceptual gap. Conversely, in the case of the ventilator, the individual concepts considered to jointly express the phenomenon at hand stand against each other in an A vs ¬A relation. This is because some of their constitutive principles are mutually incompatible. Indeed, DEATH is usually expressed in terms of the absence of life, and LIFE as a condition opposite to death. This means that, on the one hand, the intersection of the two concepts is unable to satisfactorily express the phenomenon later conceptualized as BRAIN DEATH. On the other hand, the concepts that can be used to approach (albeit only approximately) the phenomenon at hand cannot be meaningfully combined, because combining them would result in a logical inconsistency. This raises a conceptual conflict.

Conflicts, gaps, and conceptual change

We want to point out that conceptual conflicts and conceptual disruptions can occur at different times in the very same process, and this has led to some confusion in the literature. For example, Hopster (2021, p. 4) suggests that the advent of ventilator technology in the mid twentieth century caused a conceptual disruption in the form of an overlap for the concepts LIFE and DEATH, whereas Löhr (2022, p. 845) claims that the ventilator case is an example of a conceptual gap because it led to the introduction of the new concept BRAIN DEATH.

To pursue this matter, consider in more detail the invention of the mechanical ventilator and the conceptual disruption it caused, which is a popular example in the ethics of technology literature (Nickel, 2020; Nickel et al., 2022). Ventilators were invented in the 1950s and doctors realised that they could keep the basic processes of life like breathing, heart beating, cellular respiration going in patients even if those patients had no brain activity and would never again wake up or perform any of the higher functions of human life like talking or being aware of one's surroundings. This development led to a conceptual disruption in our concepts of life and death because these concepts at the time were based in part on whether a person was breathing on their own.

Here we focus on the concept DEATH, which had at the time as a constitutive principle that a person is dead if and only if that person is not breathing on their own. However, it was also standard to treat anyone whose heart was obviously beating as not dead. One problem was that, after ventilators became available, the two criteria came apart. This is a conceptual conflict since some patients now seemed to be both dead (not breathing on their own) and not dead (heart beating).

Although the history is complex (because of developments like organ transplantation, cardiopulmonary resuscitation, and the electroencephalogram) and we are idealising to some extent to highlight the features that illuminate our account of conceptual disruption, it makes sense to break down the development into the following stages. At first, we have the old conceptual scheme where DEATH is tied to *not* breathing on one's own. The new technology of ventilators then causes the *conceptual conflict* just described pertaining to the concept DEATH and the new facts about some patients on ventilators. We explained this conflict in terms of inconsistency between empirical facts (i.e., patients with no possibility of living on their own can be kept breathing, and their hearts can be kept beating, by a new technology).

Next, experts came to think that heartbeat and breathing were not the best criteria for understanding DEATH, which instead ought to be defined in terms of lack of brain activity. Thus, after this conceptual change, some patients whose hearts are beating and are on ventilators are actually dead. However, this change in the concepts LIFE and DEATH away from the lungs and heart and toward the brain left a *conceptual gap*—how do we label patients who are on ventilators, and so are breathing and whose hearts are beating, but have no brain activity? This question is especially pressing because these are the best candidates for organ transplantation since organs quickly deteriorate once the heart stops beating.

A second conceptual change then occurred when the concept of *brain death* was introduced to classify patients who fit into this category. Therefore, in the development of a new conceptual scheme where LIFE and DEATH are defined in terms of brain activity, we see *both* a conceptual conflict *and* a conceptual gap, but these occur at different times in the development. The conceptual conflict occurs first and spurs one kind of conceptual change—redefining the concepts of life and death. This first conceptual change leaves a conceptual gap and so spurs another kind of conceptual change—introducing a new concept to use on those patients that fall into the gap.

Table 2 Old conceptual scheme

Breathing on own?	Yes	No
Heart beating?	Yes	No
Concept applied	Alive	Dead

Table 3 Conceptual conflict after ventilator technology

Breathing on own?	Yes	No	No
Heart beating?	Yes	Yes	No
Concept applied	Alive	? (CONFLICT)	Dead

Table 4 New conceptual scheme with conceptual gap

Breathing on own?	Yes	No	No	No
Heart beating?	Yes	Yes	Yes	No
Brain activity?	Yes	Yes	No	No
Concept applied	Alive	Alive	Dead (GAP)	Dead

Table 5 New conceptual scheme without conceptual gap

Breathing on own?	Yes	No	No	No
Heart beating?	Yes	Yes	Yes	No
Brain activity?	Yes	Yes	No	No
Concept applied	Alive	Alive	Dead (BRAIN DEAD)	Dead

Table 6 Old conceptual scheme

Orbits sun?	Yes	Yes
Sphere?	Yes	No
Are there only a few?	Yes	No
Concept applied	Planet	Not planet

Table 7 Conceptual conflict after Kuiper belt discoveries

Orbits sun?	Yes	Yes	Yes
Sphere?	Yes	Yes	No
Are there only a few?	Yes	No	No
Concept applied	Planet	? (CONFLICT)	Not planet

Table 8 New conceptual scheme with conceptual gap

Orbits sun?	Yes	Yes	Yes
Sphere?	Yes	Yes	No
Are there only few?	Yes	No	No
Concept applied	Planet	Not planet (GAP)	Not planet

Table 9 New conceptual scheme without conceptual gap

Orbits sun?	Yes	Yes	Yes
Sphere?	Yes	Yes	No
Are there only a few?	Yes	No	No
Concept applied	Planet	Not planet (DWARF PLANET)	Not planet

The following four tables illustrate the development just described and highlight the two different kinds of conceptual disruption (conflict and gap) which occur at different times and the two kinds of conceptual change which also occur at different times in the development. The labels on the left display the shifting constitutive principles for being alive, being dead, and being brain dead (Tables 2, 3, 4, 5).

We want to emphasise that conceptual gaps and conceptual conflicts do not always display this pattern, but one can find this sort of development in a range of examples. Consider again the redefinition of PLANET discussed above.

The following tables illustrate the development described above and highlight the two different kinds of conceptual disruption (conflict and gap) which occur at different times in the development and the two kinds of conceptual change which also occur at different times in the development. The labels on the left display the shifting constitutive principles for being a planet, not being a planet, and being a dwarf planet. Again, these principles pertain to whether the object orbits the sun, whether the object is in hydrostatic equilibrium from its own gravity so as to take on roughly spherical shape, whether the object is in a class with few members, and whether the object has cleared debris from the neighbourhood of its orbit (Tables 6, 7, 8, 9).

In the case of the development of PLANET and DWARF PLANET, we see roughly the same pattern as in the example with LIFE, DEATH, and BRAIN DEATH. The old conceptual scheme is disrupted first by a conflict that occurs after the introduction of a new technology. Then a new conceptual scheme is introduced that eliminates the conflict but contains a conceptual gap. Finally, a new concept is introduced to fill the conceptual gap.

Severity of conceptual disruption

We assign different degrees of severity to different instances of conceptual disruption, based on a combination of the type of conceptual disruption involved and the level at which such disruption occurs.[12] We consider severity not merely in terms of a higher or lower degree of uncertainty, but in terms of urgency for a conceptual change.[13]

Mild conceptual disruption

We classify an instance of conceptual disruption as mild when it does not give rise to conceptual uncertainty. An example of this is a conceptual disruption occurring at the level of the conceptual scheme that manifests itself as a conceptual gap understood in terms of the non-exhaustiveness of the conceptual scheme under consideration. This type of conceptual disruption does not give rise to conceptual uncertainty, but merely uncovers the absence of a needed concept. Indeed, when discussing conceptual gaps, there is a salient difference between not knowing which concept(s) to apply and knowing that no concept applies. Just like conceptual disruption is a type of change without being a type of *conceptual* change, conceptual disruption entails uncertainty, but it does not necessarily entail *conceptual* uncertainty.[14]

[12] It should be noted that the severity attributions to follow are merely indicative and may vary depending on the concepts under consideration in a given case. Specifically, the degree of severity may vary depending (among others) on the range, depth, and valence of impacts of an instance of conceptual disruption (Hopster, 2021). For example, it seems reasonable to expect that the anticipated severity may increase when foundational concepts are being disrupted, e.g., logical concepts and moral concepts.

[13] As such, despite the use of similar terminology, this should not be confused with the discussion regarding deep and shallow accounts of conceptual disruption as covered in Löhr (2023a).

[14] Let us consider the following example. I am getting ice cream and I can choose among several different flavours. There is a salient difference in terms of the nature of my uncertainty if I cannot choose a flavour because I enjoy several flavours equally or I do not particu-

For example, a conceptual gap may result from a new technology being developed for which we do not yet have a concept. In such a scenario, we know that what we want to convey cannot be expressed using any of the concepts in our current conceptual repertoire. Therefore, it is possible for a conceptual scheme to not be exhaustive without such a conceptual gap raising questions related to the application conditions of each of the concepts included in that scheme.

Moderate conceptual disruption

We classify an instance of conceptual disruption as moderate when it gives rise to conceptual uncertainty, which can be mitigated provisionally, while a more adequate and permanent solution is being developed. An example of this is conceptual disruption occurring in the form of conceptual gaps intended as non-exhaustiveness at the level of conceptual clusters and individual concepts. For example, while the combination of "moral autonomy, human dignity, identity, personhood, liberty, anonymity and confidentiality" (van de Poel, 2020, p. 55) cannot fully express the phenomenon that can be conveyed by PRIVACY, it can provide an approximation of such a phenomenon that can serve as a meaningful placeholder while the conceptual disruption is addressed and the conceptual gap bridged.

Severe conceptual disruption

Lastly, we classify an instance of conceptual disruption as severe when it gives rise to conceptual uncertainty, which cannot be meaningfully mitigated provisionally, i.e., before it is addressed and solved. An example of this is conceptual disruption occurring in the form of conceptual conflicts at the level of conceptual clusters. For example, in the case of the ventilator, ALIVE and DEAD cannot be combined to meaningfully serve as a placeholder to express the phenomenon later conceptualized as BRAIN DEATH, due to such concepts being mutually contradicting. This means that there is no effective temporary solution to the conceptual disruption, which signals the urgency of addressing the instance of conceptual disruption. Indeed, the literature seems to confirm this hypothesis, as the disruption of ALIVE and DEAD in the case of the ventilator has been addressed urgently after the introduction of the technology, and has prompted an urgent assessment and reconceptualisation of DEATH, which ultimately led to the introduction of BRAIN DEATH (Baker, 2013).

Footnote 14 (continued)

larly enjoy any flavour, and if I cannot choose a flavour because every flavour contains milk and I am lactose intolerant. In the first case, I am comparing flavours that I deem to be (at least) somewhat suitable choices, and my uncertainty lies in trying to understand which flavour I would prefer in that specific instance. I do not know which works best. In the second case, I know that no flavour of ice cream would be a suitable choice. I know that nothing works.

Comparison

Overall, we have offered many suggestions for how to develop the idea of a conceptual disruption, and many of these build on the existing literature. *First*, we suggest that conceptual disruptions be characterised in terms of interruptions in the normal functioning of a concept, cluster of concepts, or conceptual scheme. The definition we propose differs from previous suggestions in ways outlined in "Background on conceptual disruption". *Second*, we offer a new terminology for a kind of conceptual disruption. Instead of calling them "overlaps", we use "conflicts", which brings out the sense in which the concepts involved indicate conflicting or incompatible information. *Third*, we explain conceptual conflicts in terms of the constitutive principles of the concepts involved. This novel explanation fits well with our account of conceptual disruption, while also representing a significant step forward in the literature investigating the foundations of the phenomenon of conceptual disruption. *Fourth*, we offer a substantive account of the difference between conceptual disruption and conceptual change. Conceptual disruption is an interruption in the functioning of concepts, whereas conceptual change occurs when there is a change in the constitutive principles for the concepts involved. A difference in constitutive principles is a difference in concepts. Although it can, conceptual disruption need not force any conceptual changes. This account builds on the way these terms have been used in the literature on conceptual disruption and the literature on conceptual engineering. *Fifth*, we emphasise the distinction between disruptions involving a single concept, multiple concepts, and entire conceptual schemes. This distinction is essential for correctly characterising conceptual gaps, which are the other main kind of conceptual disruption we consider. This emphasis differs from what is found in the literature on conceptual disruptions. *Sixth*, we incorporate this distinction with the distinction in severity of conceptual disruptions, intended in terms of urgency for a conceptual change. Together, these distinctions allow for a fine-grained classification system of conceptual disruptions that builds on those found in the literature. *Seventh*, we explained how distinct kinds of conceptual disruptions can occur together in a single event. We provided two case studies, one involving the artificial respirator and one involving the definition of *PLANET*. In each case, there is a conceptual conflict caused by some new technology, which is then followed by a conceptual gap caused by how people reacted to the conceptual conflict. This dynamic relationship between kinds of conceptual disruptions has not been noticed before, and

our account consistently explains why some theorists have labelled these cases as gaps and others have labelled them as conflicts. Moreover, we find the same pattern in two disparate examples. *Eighth*, our account of conceptual disruption fits well with existing accounts of kinds of conceptual engineering (explained in the next section). Explaining how our conceptual disruption in general, and various kinds in particular, interact with the process of conceptual engineering is a major benefit of our view. Although the connection between the two areas of research is self-evident and has been noted in the literature on conceptual disruptions, our insight about how certain kinds of conceptual engineering are appropriate for certain kinds of conceptual disruptions is novel and significant.

The role of conceptual engineering

We have discussed the literature on conceptual disruption, provided new accounts of conceptual conflict and conceptual gaps, and applied these accounts to some examples of conceptual disruption. In each of the examples, it is clear that conceptual engineering can be a response to conceptual disruptions. *Conceptual engineering* is the topic of how to assess our concepts and how to improve concepts that turn out to be problematic (Cappelen & Plunkett, 2020). In each of the cases discussed in the previous section, there were multiple cases of conceptual engineering corresponding to the multiple cases of conceptual disruption. In each case, a new technology led to a conceptual conflict, and this conceptual conflict was eliminated by redefining the old concepts in a way that allowed them to be applied to new situations brought about by new technologies. In the ventilator case, the concepts of life and death were redefined, while in the planet case, the concept of planet was redefined. The redefinition eliminated the conceptual conflict in each case but led to a new conceptual disruption—a conceptual gap. In the ventilator case, the gap pertained to how to label patients on ventilators with no brain activity. In the planet case, the gap pertained to how to label celestial objects that had not cleared their orbits of debris. Then, in each case, a new concept was introduced to fill the gap. In the ventilator case, the new concept is BRAIN DEATH, and in the planet case, the new concept is DWARF PLANET.

In his recent influential discussion of conceptual engineering, Chalmers (2020) distinguishes two kinds of conceptual engineering: *re-engineering*, where existing concepts are updated or replaced and *de novo engineering*, where wholly new concepts are introduced. We want to draw attention to how this distinction shows up in our discussion of the two examples. In each case, the development begins with a new technology that causes a conceptual conflict, and then that conceptual conflict is removed by re-engineering. In these cases, the old concepts that display a problematic conflict are re-engineered with new concepts that are meant to do roughly the same job but without giving rise to the problematic conflict. Moreover, in each case, the re-engineering to avoid conceptual conflict causes a new conceptual disruption in the form of a conceptual gap, which is then removed by de novo engineering. In these cases, an entirely new concept is introduced to classify phenomena that fell into the conceptual gap. Thus, it seems like re-engineering is the response to a conceptual conflict, whereas de novo engineering is the response to a conceptual gap. Moreover, the conceptual gap is caused by the re-engineering effort in light of the next technology in question. We want to emphasise that this pattern need not hold for all examples, but it does seem reasonable to see this distinction between kinds of conceptual engineering as a result of the distinction between kinds of conceptual disruption.

Some conceptual disruptions do not call for re-engineering or de novo engineering; instead, eliminating the concepts in question is the appropriate course of action. These conceptual disruptions are in the form of conflicts, and take place at the second and third level of conceptual disruption, affecting conceptual clusters and individual concepts. Such instances of conceptual disruption do not raise uncertainty about the constitutive principles or application conditions of the concept(s) under consideration, but uncover an incompatibility between one or more concepts and some facts about the world. Indeed, there is a salient difference between *not knowing* the constitutive principles of a concept or *not knowing* when and how a concept can be applied, and *disagreeing* with the application conditions of a concept.[15] In these cases, such conceptual disruptions highlight grounds for the discontinuation or abandonment of one or multiple concepts. An example of this type of disruption can be drawn by considering how technology can cause some of our concepts to no longer apply, thus prompting the abandonment of such concepts.

Conclusion

New technologies have the power to disrupt our social structures and even the very concepts we use to represent and understand the world. In this paper, we have critically engaged with recent work on conceptual disruption and

[15] For example, when an instance of conceptual opposition concerns a derogatory concept (e.g., a concept imbued with racial, sexist, homophobic, or ableist undertones), this does not mean that we do not understand the constitutive principles of such a concept, or that we do not understand how and when the concept can be applied. Rather, the instance of conceptual opposition uncovers a rupture such that we do not think the concept *should* be applied.

offered an alternative account of the phenomenon of conceptual disruption, not only in terms of a new way to divide the kinds of conceptual disruption, but also as substative accounts of the general category of conceptual disruption and two of its types: gaps and conflicts. We appeal to the idea that concepts have functions in order to explain *conceptual disruption* in general—i.e., interruptions in the normal functioning of a concept, cluster of concepts, or conceptual scheme. We distinguish conceptual disruption from *conceptual change* because the latter involves the introduction or elimination of a concept or the changing of a concept's constitutive principles. A *conceptual gap* occurs when there is some phenomenon that ought to be classified at a certain level of precision but no relevant concepts properly apply to the phenomenon in question. A *conceptual conflict* occurs just in case there is some known fact that is inconsistent with some constitutive principles for some concepts. We argue that it makes more sense to classify certain conceptual disruptions as conceptual conflicts rather than the established category of conceptual overlaps.

In addition, we explain how conflicts and gaps occur at different times in processes by which new technologies lead to changes in our concepts. In particular, we explained how gaps and conflicts show up in both the case of the mechanical ventilator (which spurred conceptual changes leading to the new concept of brain death) and astronomical technologies (which spurred conceptual changes leading to the new concept of dwarf planet). In each case, we saw that the new technology led to a conceptual conflict, that was followed by a conceptual change. This conceptual change introduced a conceptual gap, which was then addressed by introducing a new concept to classify the phenomena in question. Throughout our study of conceptual disruption, we emphasised its heterogeneous nature and the fact that it can show up at different levels, including individual concepts, clusters of concepts, and conceptual schemes. Moreover, conceptual disruptions differ in severity, which we explain by appealing to how urgent it is to address them. Finally, we defended the idea that certain kinds of conceptual engineering are appropriate for addressing different sorts of conceptual disruptions. In particular, conceptual conflicts often induce either re-engineering or abandonment, whereas conceptual gaps often spur de novo engineering where entirely new concepts are introduced.

Acknowledgements An earlier version of this paper was originally presented at the Ethics of Socially Disruptive Technologies (ESDiT) Foundation & Synthesis Research Colloquium, held April 20, 2022. We would like to express our gratitude to the participants for many helpful questions and comments. Many thanks also to the editors and to an anonymous referee for helpful comments.

Author contributions This manuscript is a collective effort that cannot be attributed to either author individually. SM took the lead on the taxonomy. KS took the lead on the constitutive principles. Both authors contributed to the writing and editing, and approved the final draft of the manuscript.

Funding This work is part of the research programme Ethics of Socially Disruptive Technologies, which is funded through the Gravitation programme of the Dutch Ministry of Education, Culture, and Science and the Netherlands Organisation for Scientific Research under Grant number 024.004.031.

References

Babushkina, D. (2020). Robots to blame. In M. Nørskov, J. Seibt, & O. S. Quick (Eds.), *Culturally sustainable social robotics: Proceedings of Robophilosophy* (pp. 305–315). IOS Press.

Baker, R. (2013). *Before bioethics*. Oxford University Press.

Brożek, B., & Janik, B. (2019). Can artificial intelligences be moral agents? *New Ideas in Psychology, 54*, 101–106. https://doi.org/10.1016/j.newideapsych.2018.12.002

Burke, B. (1992). Searching for exoplanets. *Targets for space-based interferometry* (pp. 81–83). European Space Agency.

Cappelen, H. (2018). *Fixing language*. Oxford University Press.

Cappelen, H., & Plunkett, D. (2020). A guided tour of conceptual engineering and conceptual ethics. In A. Burgess, H. Cappelen, & D. Plunkett (Eds.), *Conceptual engineering and conceptual ethics* (pp. 1–26). Oxford University Press.

Carey, S. (2009). *The origin of concepts*. Oxford University Press.

Cartlidge, E. (2012). Loose cable may unravel faster-than-light result. *Science, 335*(6072), 1027. https://doi.org/10.1126/science.335.6072.1027

Chalmers, D. J. (2020). What is conceptual engineering and what should it be? *Inquiry*. https://doi.org/10.1080/0020174x.2020.1817141

DeCew, J. (2018). Privacy. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Stanford University.

Eklund, M. (2002). Inconsistent languages. *Philosophy and Phenomenological Research, 64*(2), 251–275. https://doi.org/10.1111/j.1933-1592.2002.tb00001.x

Hopster, J. (2021). What are socially disruptive technologies? *Technology in Society, 67*, 101750. https://doi.org/10.1016/j.techsoc.2021.101750

Hopster, J., Brey, P., Klenk, M. B. O. T., Löhr, G., Marchiori, S., Lundgren, B., & Scharp, K. (2023). Conceptual disruption and the ethics of technology. In I. van de Poel, L. Frank, J. Hermann, J. Hopster, D. Lenzi, S. Nyholm, B. Taebi, & E. Ziliotti (Eds.), *Ethics of Socially Disruptive Technologies: An Introduction* (pp. 141–162). Open Book Publishers. https://doi.org/10.11647/obp.0366.06

Hopster, J., & Löhr, G. (2023). Conceptual engineering and philosophy of technology: Amelioration or adaptation? *Philosophy and Technology*. https://doi.org/10.1007/s13347-023-00670-3

International Astronomical Union. (2006). *Definition of a planet in the solar system*. IAU Resolution B5.

Jorem, S. (2022). The good, the bad, and the insignificant. *Synthese, 200*, 1–20. https://doi.org/10.1007/s11229-022-03548-7

Löhr, G. (2022). Linguistic interventions and the ethics of conceptual disruption. *Ethical Theory and Moral Practice, 25*, 835–849. https://doi.org/10.1007/s10677-022-10321-9

Löhr, G. (2023a). Conceptual disruption and 21st century technologies: A framework. *Technology in Society, 74*, 102327. https://doi.org/10.1016/j.techsoc.2023.102327

Löhr, G. (2023b). Do socially disruptive technologies really change our concepts or just our conceptions? *Technology in Society, 72*, 102160. https://doi.org/10.1016/j.techsoc.2022.102160

Löhr, G., & Michel, C. (2023). Conceptual engineering, predictive processing, and a new implementation problem. *Mind & Language*. https://doi.org/10.1111/mila.12471

Machery, E. (2009). *Doing without concepts*. Oxford University Press.

Nickel, P. J. (2020). Disruptive innovation and moral uncertainty. *NanoEthics, 14*(3), 259–269. https://doi.org/10.1007/s11569-020-00375-3

Nickel, P. J., Kudina, O., & van de Poel, I. (2022). Moral uncertainty in technomoral change: Bridging the explanatory gap. *Perspectives on Science, 30*(2), 260–283. https://doi.org/10.1162/posc_a_00414

Putnam, H. (1965). A philosopher looks at quantum mechanics. In R. G. Colodny (Ed.), *Beyond the edge of certainty* (pp. 75–101). Prentice-Hall.

Riggs, J. (2021). Deflating the functional turn in conceptual engineering. *Synthese, 199*, 11555–11586. https://doi.org/10.1007/s11229-021-03302-5

Santoni de Sio, F., & Mecacci, G. (2021). Four responsibility gaps with artificial intelligence: Why they matter and how to address them. *Philosophy & Technology, 34*, 1057–1084. https://doi.org/10.1007/s13347-021-00450-x

Scharp, K. (2013). *Replacing truth*. Oxford University Press.

Sebastián, M. Á. (2021). First-person representations and responsible agency in AI. *Synthese, 199*(3–4), 7061–7079. https://doi.org/10.1007/s11229-021-03105-8

Sommaggio, P., & Marchiori, S. (2018). Break the chains: a new way to consider machines' moral problems. *Biolaw Journal, 5*(3), 241–257. https://doi.org/10.15168/2284-4503-339

Sullins, J. P. (2009). Artificial moral agency in technoethics. In R. Luppicini & R. Adell (Eds.), *Handbook of research on technoethics* (pp. 205–221). IGI Global. https://doi.org/10.4018/978-1-60566-022-6.ch014

Thagard, P. (1993). *Conceptual revolutions*. Princeton University Press.

van de Poel, I. (2020). Core values and value conflicts in cybersecurity: Beyond privacy versus security. In M. Christen, B. Gordijn, & M. Loi (Eds.), *The ethics of cybersecurity: The international library of ethics, law and technology* (Vol. 21, pp. 45–71). Springer. https://doi.org/10.1007/978-3-030-29053-5_3

van de Poel, I., Frank, L., Hermann, J., Hopster, J., Lenzi, D., Nyholm, S., Taebi, B., & Ziliotti, E. (Eds.). (2023). Ethics of Socially Disruptive Technologies: An Introduction. Open Book Publishers.

Warren, B., & Brandeis, L. (1890). The right to privacy. *Harvard Law Review, 4*(5), 193–220.

Zeng, Y., Gordiichuk, P., Ichihara, T., Zhang, G., Sandoz-Rosado, E., Wetzel, E. D., Tresback, J., Yang, J., Kozawa, D., Yang, Z., Kuehne, M., Quien, M., Yuan, Z., Gong, X., He, G., Lundberg, D. J., Liu, P., Liu, A. T., & Strano, M. S. (2022). Irreversible synthesis of an ultrastrong twodimensional polymeric material. *Nature, 602*(7895), 91–95.