NECSUS

Clipping us together: The case of the Google Clips camera

Oshri Bar-Gil

Abstract
This article uses the Google Clips camera as a case study to illustrate the impact of autonomous machine learning on self-perception, and to investigate how 'delegation' of our self to those cameras occurs. The research is based on reviews of the Google Clips camera, analysed using Computer Assisted Qualitative Data Analysis (CAQDAS) and interpreted using Don Ihde's postphenomonological framework complemented by Bruno Latour's relation analysis. Positioning the Clips camera as a technological mediator, the analysis concentrates on human-technology-world interaction relations. The research findings include changes in self-perception through complex concepts, such as autonomy, agency, and rationality.

Keywords: AI, artificial intelligence, computer vision, google, Google self, intelligence, self-concept

Introduction

What if we could build a product that helped us be more in-the-moment with the people we care about? What if we could actually be in the photos, instead of always behind the camera? What if we could go back in time and take the photographs we would have taken, without having had to stop, take out a phone, swipe open the camera, compose the shot, and disrupt the moment? And, what if we could have a photographer by our side to capture more of those authentic and genuine moments of life...[1]

The Google Clips camera is an intriguing and tempting prospect. Who would not want to use an 'autonomous' camera, boasting cutting-edge technology

paired with machine learning algorithms, in order to be more 'in the moment'? Who would not want to be in the picture, rather than behind the camera; to possess a 'photographer by his side to capture more of those authentic and genuine moments of life'? And achieving all of this without having to think about shutter configuration, composition, and the objects in the photo.

Functioning as a mediator, the autonomous Clips camera is a part Google's drive to create and maintain a digital Doppelgänger. Users delegate intentions, actions, and choices to their doppelgänger – enhancing, but also possibly undermining, their autonomy as human subjects. The Clips camera is just one of many technology consumer lifestyle products produced by Google. When considered as part of an interlinked network of products and services offered by Google and other platforms, it is possible to identify the beginnings of a trend, pointing toward the erosion of our autonomous thought and action.

In the section that follows, I describe the Google Clips camera and its place in the field of digital photography, before describing the theoretical background to the study. The next section outlines the research methodology, followed by the results of the analysis and a discussion of the findings. I then conclude with a brief consideration of the implications of my findings, and potential areas for future research.

Google Clips

In October 2017, after a three-year development process, Google, the multinational technology company, unveiled a new hardware product, Clips. A two-inch compact camera which could be 'clipped' to clothing or other objects (thus the name), the Clips camera was distinguished by its autonomous operation. The camera's artificial intelligence software architecture enabled it to independently determine when to take a picture, guided by its algorithmic characterisation of a specific event as 'interesting'.[2] The selection of such events is facilitated by the camera's algorithm 'learning' to identify significant associations connected to the camera's user, such as family members, pets, locations, and events. This learning guides the camera in generating 'suggested clips', which the user is more likely to save.[3] A key feature of Clips is its independent capacity to determine the best timing for a photograph, and then self-processing the image by means of its hardware and algorithms – without the need for human intervention.[4]

In late 2019, Google decided to withdraw the Clips camera from the consumer market.[5] Unlike other discontinued Google hardware products, such as Google Glass, this was not due to concerns about privacy intrusions, but something more mundane: 'the device was panned in reviews for poor image quality and for not recording sound for its video clips'.[6] Comments and feedback included complaints by early users that the camera was 'too unpredictable'[7] for their liking.

Digital photography is not a new phenomenon. The first digital cameras became available on the mass consumer market in the late 1990s. Google was not a pioneer in developing the digital photography consumer market. However, the synchrony of Google's suite of technology products, together with their accessibility (most operate on a free or 'freemium' model) has underpinned the evolution of what can be described as 'digital memory': the ability to store virtually unlimited numbers of images, together with the capacity to categorise, analyse, retrieve, and edit this data, with or without direct human intervention. The reduced cost of on- and offline storage media, in conjunction with the smartphone revolution – in effect, placing a digital camera in the pocket of billions of people – has been key to this evolution, through the image data analysis that makes innovations like the Google Clips camera possible. It is estimated that at the present time, we shoot (and, by default, save) the staggering amount of 20.5 billion images a day.[8]

Digital photography, and the use of digital photography platforms as a repository of events, social connections, and memories, is ubiquitous. Between this and the sheer volume of digital images produced on a daily basis, it seems surprising that 'unpredictability', in and of itself, caused the failure of Google Clips. This paper seeks to explore the market experiment (and failure) from another perspective: its potential to function as a reliable photographic self-doppelgänger. I am interested in exploring the experiences of early-adopters of the Clips camera; specifically, how might the use of an autonomous camera as a mediator of lived experiences change one's perception of self?

The self and self-concept

The meaning of 'self' is continuously shaped and altered by changes in culture, the social sphere, and in technology. With regards to this article, I am

interested in the changes for our *self-concept*. I will cite Baumeister as a working definition: 'The individual's belief about himself or herself, including the person's attributes and who and what the self is.'[9] In his view, the self-concept emerges from social and linguistic meaning. The changes in temporal and spatial notions of self have shaped sociological discussions about the self in the modern age.[10] The technology of the late information age has the potential to change how we consider these questions,[11] particularly with respect to the sense of self autonomy.[12] In considering the notion of autonomy, I rely on Searle and his predecessors' notion of autonomy as influenced by his definition of intention, as 'the psychological states that produce and guide action'.[13]

From philosophy of technology to technological mediation

An influential school of contemporary philosophers and thinkers about technology, including Latour, Ihde, and Mitcham,[14] have critiqued the pessimistic and essentially dystopian assessment of technology presented by an earlier cohort of thinkers, including Ellul, Heidegger, and Jaspers. This latter school of thought argues that their predecessors sought to impose a uniform model of attitudes toward 'Technology'. In opposition to this, the new generation of technology thinkers have proposed new paradigms for conceptualising the relationship between technology and lived experience, focusing principally on the role of technology as a mediator, building and shaping the relationship between human users and the wider environment; thinking of technology acting as a catalyst, enzyme, or inert linkage, depending on use, context and meaning.[15]

Ihde's taxonomy of the differing experiences of technology identifies four types of mediating relationships. The first, *embodiment*, describes the experience of technology as though it were a part of our bodies. Ihde's second category is the *hermeneutic*, where technology allows one to 'read' the world better. One example is digital thermometers such as a weather app, which help us 'feel' climatic conditions in other physical locations. The third category is *alterity*. This category conceptualises 'technology' as an independent object or physical presence in and of itself. A good example of alterity relations would be humanoid robots, capable of some form of two-way communication with their human users.

Ihde's final category is the *background* relation. In this form, technology operates in the background of everyday life. It does not occupy an overt place in our everyday lives; but nevertheless, it constitutes an integral part of a landscape saturated with countless technologies.[16] Collectively, what we can take from this is the notion of co-construction, and how this extends the mediating relationship beyond the human user and into the world.[17]Latour, for his part, based his analysis of the relationship between humans and technology in terms of 'actor' and 'network'. Consequently, his conceptualisation is often described as 'actor-network theory' (ANT).[18] Following Latour, if one thinks symmetrically, then one must conclude that agency cannot be restricted to human beings – which, incidentally, explains his preference for the term 'actants', rather than 'actors'.[19]

Latour explicates this distinction through the concept of technical mediation, which he describes in a number of ways. One example is that of *translation*. When a technological tool mediates a relationship, it does so through the 'translation' of a 'program of action'. Another example, according to Latour, is *composition*. Composition assumes that mediation incorporates not just the translation of programs of action, but also (and simultaneously) the linkage of actants into action.[20]

This paper focused on *delegation*, which is Latour's principal category of technology mediation. Latour famously used the example of a speed bump to capture the notion of delegation. The desired program of action (for drivers, to reduce their speed) is 'inscribed' in asphalt, performing a function not dissimilar from that of a policeman signaling speeding drivers to slow down.[21]

Questions relating to surveillance and privacy are a central aspect of the broader discourse about technology in general, and Google's increasing presence as a panopticon-like entity in particular. This consideration of the Google Clips camera will not explore these issues in detail, however, for two reasons. First, there are numerous, comprehensive assessments of these issues in relation to the use of autonomous products and platforms.[22] That aside, a fundamental aspect of research into changes in self-perception from a phenomenological perspective is the focus on the actual experience of users vis-a-vis the product. Given that neither privacy nor surveillance were mentioned as key factors by the users surveyed in this study, it seems redundant to explore this particular issue further in the present context.

Methodology

Any attempt to analyse the influence of autonomous products and services on self-perception must engage with two fundamental issues: the wide range of products and services offered to consumers, and the number of users of these products. This is particularly pertinent in an analysis of a product made by Google, the largest technology company in the world.

With regards the first point, Google is made up of a large suite of inter-connected services. Whilst remarkably diverse in terms of user functionality and accessibility, they collectively constitute a common – and extremely large – pool of data. In relation to the second, some estimates suggest that Google has as many as 2.2 billion users.[23] Given the breadth of data available, it was necessary to delineate the scope and focus of the research data, to ensure its direct relevance to the research question.

I chose content analysis as my key research tool in order to steer my analysis away from an abstract and ungrounded discussion about autonomous products. Specifically, I used a special type of content analysis: netnography.[24] This method allows for the collection of data in naturalistic settings with minimal intervention from the researcher, thus creating a range of possibilities for subsequent analysis.[25]

Netnography studies cultural phenomena by analysing online discourse in social networks and communities. Researching a user community of 2.2 billion users is, self-evidently, a daunting task, and thus requires some form of delineation. I settled on the community of technology reviewers for the consumer media, by definition 'early adopters' of emergent technologies. As a technologically literate group of users, keen to try products that do not have existing analogies in the technology consumer market, it is reasonable to assume that they would be influenced by such use more than subsequent users. Of the several forms of collating data for netnography, I chose to use blog research, concentrating on five key technology publications that operate technology review blogs: *The Verge*, *Wired*, *Engadget*, *Ars-Technica*, and *The Keyword* (the last is maintained by Google).[26] These publications were selected due to their reputation as leading representational techno-cultural publications.

All five blogs were mined for references to the Clips camera; data mining was carried out manually. In each publication, a search for reviews and comments relating to the Google Clips camera was conducted, and the webpage search results were saved as PDFs for further analysis. Typically, the first blog

reviews are published at the time of a product's launch. While the product is not commercially available at this time, and as such would not have attained its full societal impact, this time point is optimal for identifying the changes in routines and usage that ensue from the introduction of new technology. While mining from the various sites, I focused on articles and responses that specified how users experienced the world differently while using the product.

Atlas.ti, a computer-assisted qualitative data analysis software programme, was used to collect and code all the data mined.[27] The coding scheme was built from grounded codes and theoretical post phenomenological concepts mentioned earlier. The analysis and interpretation of the data was guided by Latour and Ihde's postphenomonology theoretical framework and by Searle's concept of autonomy and intention, which I will discuss later.[28]

How did we get here? Creeping towards the Clips camera

Expanding from Gibson's affordances theory, which adopts the term 'affordance' to describe the different 'possibilities for action' provided by different technologies,[29] it is possible to identify several possible actions created by the phenomenon of digital photography, as compared to its physical counterpart: editing, manipulating, storing, retrieving, and so on. One aspect of the next phase of the information age, commonly described as the Fourth Industrial Revolution,[30] relates to the possibilities afforded by the exponential leap in the availability of data volume and cloud storage, coupled with developments in software algorithms, high networking rates, and hardware capabilities. The underlying presumption is that the synergy enabled by these changes will have a paradigm-shifting impact on how we live. With regards to photography specifically, the ready availability of technological capabilities in the late digital age point to one nexus – the shrinkage, by several magnitudes of order, in the marginal cost of storing and accessing digital photographs. This change has completely reconfigured the underlying purpose and potentiality of photography.

Nevertheless, physical interaction between user and object is still required in order to 'perform' photography; the human photographer requires a digital camera of some sort, the technical knowledge required to operate it, and the spatial and social awareness necessary to capture an image as and when

desired. Human intervention, by way of decision making, is still required for identifying, composing, and capturing a specific image – 'point and shoot' requiring somewhat more in terms of individual agency than the catchphrase suggests. The unique selling point of the Clips camera, when it was introduced to the consumer market, was that it made it possible for a user to 'delegate' the intention of photography: the act, the many decisions that precede the act, and the subsequent creation of a digital memory. Over and above the creation of a digital memory, digital photography also incorporates a number of subsidiary affordances, by virtue of the metadata that is created alongside a digital image – location, time, participants, and so on.

Given Google's digital dominance it is no surprise that it is the preeminent location for saving and sharing these digital memories – through Gmail, Google Drive, and so on. These manifold usages grant Google, directly and indirectly, access to an almost infinite resource of user-generated visual data, which it can use in training its machine learning algorithms to accomplish a range of tasks hitherto unachievable. The attraction, from the archetypal user's perspective, is that this facilitates the creation of a digital doppelgänger – tacit or explicit acknowledgment of changing aspects of one's notion of self, and self-perception.

The creation of a digital doppelgänger unfolds in a number of complementary ways. First, Google is not a person, but rather a virtual entity; its very un-corporeality may afford users the emotional or intellectual distance that otherwise may inhibit such sharing behaviour, and indeed may strengthen their comfort and confidence in using the platform. Beyond this, Google's powerful algorithms have the capacity to enable automation of memory functions. Consider, as an example, the functioning of the Google Photos application in creating albums and event collages for its users. The application sends email alerts about the compilation of a new album; it makes automatic corrections to photos; edits short videos; and even curates special collections of photographs, featuring people whom its algorithms have identified as relevant for us. Through these manifold actions, the Google photo application personalises the digital image memories of its users, in the process transforming the user's relationship with Google to something akin to a background-mediated relationship.[31] Many of the functions that users once had to do themselves to personalise digital images – assuming that they possessed the requisite technical knowledge – are now performed automatically. This facilitates the seamless delegation of hitherto agentic actions to a digital doppelgänger.

What allows for this seamless transition? One possibility is Google's remarkable capacity for organising and recalling information – the platform helps users to 'find a needle in the haystack of the best images'.[32] Research into similar products[33] indicates that this is a two-step process, involving both machine learning and human labor.[34] The first step is human agency, users actively tagging photographs, designating specific images/locations/people as 'favourites', deleting unwanted images, sharing images with other users, and so on; a process of organisation and taxonomy, in essence. Other key classification activities are facilitated passively. These include geotagging, shooting time, image capture data (focal length, aperture key, ISO), and camera data, which are generated automatically when the image is created. Thus, the digital doppelgänger begins to learn its human user's preferences – when and where they use their cameras, the images they like, and what they tend to do with different types of images. From this, the doppelgänger is able to construct a predictive profile, enabling it to perform these actions on the user's behalf in the future.

Some may consider it ironic that, having taught (albeit passively) the algorithm to identify personal preferences through the wealth of data created and stored on Google's servers, the process of creating a digital image is then turned on its head. But the converse is perhaps a more convincing consideration. Given its capacity to categorise data and conceptualise patterns, it might be that the logical next step would be to outsource the active stage of creating a digital image to a non-human entity – an autonomous camera. A camera that will be present at every moment of their life; that can decide when to take a photo, or of whom, on behalf of a human agent, and with a precision as accurate as that of the human. Going back to Ihde's conceptualisation, one can think of the Clips camera in terms of a movement from an alterity relationship to a background relationship; from conscious and deliberate effort in relation to the object (the camera) to achieve an outcome to accepting this process as a given, facilitated by technology. The technology is working behind the scenes, as it were, without drawing deliberate attention to itself and without the need to be activated.[35]

Delegating vs outsourcing our self-functions?

The Google Clips camera presented a unique opportunity to observe and conceptualise the delegation of individual agency to a technology platform –

in this case the multi-layered infrastructure that links the full suite of Google products. Following Latour, delegation is a form of technological mediation carried out by a combination of mediating factors that allow the player to act.[36] In information age platforms, delegation occurs in relation to central self-functions and is uniquely operated, almost the same as the outsourcing of the self.

The sense of excess information – the desire to be present in the virtual sphere, on social networks, for example – means that users must contend with an overflow of activities that require attention and effort to process. In the case of the Clips camera and digital photography, delegation is initiated and performed through processing and uploading meaningful digital images, aspects of one's digital memory, to a repository where they can be accessed at a later date. Generally, given the subjective and intimate nature of such processes, the expectation would be for personal intervention in this process – agentic action. But on the other hand, users generally desire a frictionless life, without the encumbrances of uncertainty. This desire explains the attraction of delegation – automating and outsourcing the execution of straightforward or routine tasks to a doppelgänger, capable of performing these tasks just as they would have themselves. Clips, and the Google technology that 'afforded' its functionality, facilitates this act of delegation – the 'personalisation' of the products and services of a specific platform.

Advances in information technology – volume and accessibility of data, the increased sophistication of algorithmic functions, networking, cloud computing, and hardware[37] – all create affordances for this form of delegation.[38] In the present case, the Clips camera deploys advanced machine learning algorithms for the purposes of image recognition and image processing; these facilitate the personalisation of the act of photography, by identifying the factors that characterise the digital images captured by a user. Who are the people who matter to us? What situations and events are important to us? What style of photography do we like more? Which locations and settings resonate with us more strongly than others?

In this context the two-step process described above, of human interventions and machine learning, is both complementary and self-reinforcing. The user has the incentive to invest time and effort in 'training' the algorithmic process, once convinced that ultimately this will save time – and crucially, function as though the user were directing the operation personally. As the user develops trust in the system, it then becomes easier to delegate

wider functions to the non-human agent; by receiving information and receiving wider autonomous 'user' privileges and permissions, the non-human actor reinforces its capabilities even further. The transition to a background relation, as conceptualised by Ihde, is the product of delegation turned to outsourcing. It is important to distinguish this, however, from the simple outsourcing of mechanical or basic cognitive activities, as done by earlier technologies. What occurs here is the outsourcing of capabilities that were once the exclusive province of the self: thinking, control, decision-making processes.

This poses an interesting question. If these outsourced processes are ultimately performed without our conscious knowledge or awareness, can they still be conceptualised as delegation, in the manner Latour suggests? Once a user has bestowed upon their digital doppelgänger the ability to discern and perform one's intentions, can the act still be regarded as having been mediated by virtue of the human actor 'translating' their intent?[39] And from this, would an autonomous camera remain a non-human actor, or could it possibly be viewed as an extension of the self?[40] My contention is that this unique pattern of delegation is not explained by existing theories of technological mediation, or by the extended mind theory.[41]

Can we delegate our intentions?

To understand the role of human agency in the process, we must first clarify the concept of intentionality. Intentionality, a key concept in the philosophy of self, has been defined and interpreted in many different ways over the years.[42] For the purposes of the current study, I rely primarily on John Searle's exposition of the concept of intention. Searle defines intentionality as: 'that feature of certain mental states and events that consists in their being directed at, being about of, or representing certain other entities and state of affairs'.[43] Building on Searle, Bratman considers intention – as opposed to urge – as a means by which the self regulates interactions and long-term activity, and specifically the regulation of self-activity in both the synchronic and the diachronic planes.[44] Smith, going further, defines intentions as 'the psychological states that produce and guide action'. Smith continues: 'The efficient causes of intended actions, are what rationalize actions, are what promote both the sort of intra-agential, cross-temporal coherence that allows

people to take complicated actions over time, and facilitate the sort of trans-agential coherence that allows people to act in concert with one another.'[45]

From these definitions, one may infer that outsourcing the act of photography – the creation of a digital memory – to the Google Clips camera requires, at the least, the partial delegation of a certain mental state. Once, the person taking a snapshot would have needed to have been in a specific mental state – we can call this photographic intent; now, once this state has been delegated to the Clips camera, photographic intent will be actualised at some point, initiated by the camera and on behalf of the human agent. Through this delegation of photographic intent, a user creates a continuity of action over time, by placing the autonomous camera in a space where photography can be initiated. Through the personalisation of the camera according to the user's preferences, it will take pictures on behalf of the user, and in accordance with these preferences.

Searle[46] argues that there are two different types of intention, which should be distinguished from one another. The first is intention directed towards triggering action. An example is formulating the intention to raise one's hand in 30 seconds' time. Searle calls this 'prior intention'. The second type of intention relates to the action itself – when, after the 30 seconds have elapsed, the actor indeed raises their hand. Searle calls this 'intention in action'. Searle's argument is that in every action we perform we take for granted the social context that we are embedded in. This context consists of beliefs, abilities, and possibilities as manifested in the prior intention; nevertheless, Searle indicates that action cannot be created without intention in action.

This distinction made by Searle and others[47] is important when comparing the operation of an ordinary camera to that of the autonomous Clips camera. With a smartphone or other 'ordinary' camera, for example, both prior intention and intention in action are required to initiate the process – deciding to take a picture, selecting the object(s), pointing the camera and so on. The Clip camera, however, allows the user to delegate the intention of photography – intention in action – by 'training' it to recognise the user's patterns of prior intention. Through this, the camera is able to crystallise the user's prior intentions – who to shoot, photograph composition, and so on – and then actualise these with the proper timing:'Google's definitely onto something here. The idea is an admirable first step toward a new kind of camera that doesn't get between me and my kids.'[48]

The camera produces new affordances, neither in the human nor the camera but somewhere in between, in their common space.[49] These affordances create a hardware object, a camera, able to operate just like the owner-user. The possibility of delegating one's intentions to a camera can produce new intentions in place of one's own – to photograph the children in one's place, for example. This, further, creates an interesting paradox. Some users claimed that the camera allowed them to fulfil deeper intentions, such as spending more time with the children and less with the devices: 'Clips is letting you spend more time interacting with your kids directly, without having a phone or camera separating you, while still getting some photos.'[50]

We can see, through the example of the Clips camera, the affordances of delegating certain aspects through the technological mediation of 'prior intention'. Intention is actualised by a non-human agent whom the user trusts to act as they would – encouraging the user to invest the effort necessary to this end, and thus to be able to delegate his intention to it. Pragmatically, prior intention becomes detached intention in action. Intention in action is no longer an internal mental state, but rather now exists as a 'mixed' intention – not just human and requiring a 'dance of agencies' between the user and the technology embedded in the Clips camera.

The user's intention is manifested in the 'teaching' of the camera to become a component of the user's self. This is to say, users teach their digital doppelgänger to behave like them and for them, and to automate processes for them (as they can be described by 'background relation'). But also, this grants the digital doppelgänger such a degree of autonomy that we do not need to articulate our exact intentions to it; indeed, the doppelgänger can create new intentions for us. By mediating one's intent, the Clips Camera thus becomes an 'intermediate agency'.[51] As an active mediator between the user and the world, this state clearly departs from the position set out in Ihde's postphenomenological theory, which links notions of intention to the capacity to establish its own agency and intentions (including foreign agency, algorithmic operations, and the 'googling' that is an inherent aspect of these actions). However, it is unable to constitute an entity possessing agency in and of itself, as conceptualised by Latour's exposition of inscription and delegation processes.[52] The capacity to mediate reality, after all, is contingent upon the intervention of the user, who has personalised the device to mirror their intentions. By delegating the intention the meaning is changed: instead of delegation as inscription this is delegation as prescription, allowing the

user to delegate a portion of his prior intention. This process allows the object of delegation to influence the delegation process, by expanding the delegation to the intentional domain.[53]

By focusing on the concept of intention as embodied in the action taken by the Google Clips camera, one can identify a shift in the concept of agency as it exists between a user and a device. The shift is somewhat complex but can be described – and indeed, it manifests – in terms of the delegation of agency to the digital doppelgänger, the 'Google self'. This pattern of delegating prior intention to the autonomous device allows for a detailed description beyond the usual references that simply describe a mixed agency as a 'dance of agencies' between the subject and the technological objects,[54] and demonstrates how exactly these dance steps take place.

From intention to rationality – from human rationality to the Google self-doppelgänger's hyper rationality

We have distinguished between prior intention on the part of a human agent, and action embodied in outward action in the world; the latter, as described above in relation to the Google Clips camera, can be delegated, and is largely fulfilled by the technological agency. The realisation of 'prior intent' by technological intention in action becomes a gap, which over time the user and the enabling technology will strive to close.

Searle called this gap the 'causality gap', proposing that it is a phenomenon that derives from the illusion of free will. That is to say, once the causes of action are embodied in the prior intention are put in place, the ability to not realise the prior intention allows us to preserve the illusion of free will – which, as we saw above, cannot be easily shaken off.[55] The question (or the illusion) of free will, and its influence on the notion of self-perception, is a key question in determining the intentional aspects of an action. In the case of the Clips camera, the user can decommission the camera as an act of free will; but since it has been commissioned to fulfil a specific purpose, this action undermines its capacity to realise an existing plan of action.

Following Searle, Bratman[56] interprets the prior intention as an internal plan of action, possessing some degree of flexibility in its precision and scope. Bratman considers the notion of prior intention in terms of an internal position which ultimately leads to action – prior intention and realisation –

which makes it possible to rationally assess its functionality. This conceptualisation also makes it possible to consider the realisation of an intention as a rational question, relating to possibilities and their actuation. Google's algorithms are operated through the digital doppelgänger that they enable, outlining the fulfilment of user intent though an optimised program, shaped and regulated by the vast repository of data which reinforces their existing logic-based authority. In this way, implementing the program through the various technologies that constitute the Google digital doppelgänger makes us more rational, and more likely to actualise our prior intentions.

Users of the Google Clips camera experienced the autonomous functioning of the camera as an accurate and effective means of crystallising prior intentions, such as photographing their children or their pets in various 'authentic' activities and settings. By implication, the most rational realisation of the prior intention would be to position the camera such as to enable it to autonomously record digital memories, as an expression of the efficient, fast, and accurate action initiated by the intention.

Picture this: you're hanging out with your kids or pets and they spontaneously do something interesting or cute that you want to capture and preserve. But by the time you've gotten your phone out and its camera opened, the moment has passed, and you've missed your opportunity to capture it.[57]

According to the rationale dictated by the algorithmic logic used to categorise the vast volume of data that we create and provide to Google, the data is directed to an 'unspecified' intention – one that is more 'real', revealed through the algorithmic truth that ascertains what the user 'really' likes: the television programmes the user watches, the searches that the user conducts on Google, as opposed to what the user *says* that they like, or what the user *claims* to be searching for on the search engine.[58] In the case of the Clips camera, this intention is shaped not by the images that the user has consciously highlighted as important to them, but rather by what they actually upload and mark as favourites.

Through this, one can see how improving the rationality of the action plan actualises the prior intentions of the user at a deeper, almost unconscious level, facilitating a creeping, almost imperceptible process of further delegation by users to their digital doppelgängers. This occurs despite the erosion of self-boundaries which can result from this actual delegation of one's intentions to the technological self, which itself may be under the control and alignment of Google's intentions.

What else can we learn from the Clips camera about our Google self-doppelgänger?Interestingly, user criticism was less about issues related to the increasing delegation of self-functions, or even the tedious process of 'teaching' the camera the user's preferences. Rather, criticism focused largely on the camera's actual performance:

> I can imagine that with more time and use, Clips could have serendipitously captured something truly special, but in a couple of weeks of testing, it didn't… I'm sure I could get better at using the Clips camera with practice – getting a better idea of its ultra-wide field of view, finding the best angles and positions for it, and so on – but I'm not convinced that the effort involved would be rewarded with great results.[59]

The camera, it seems, did not fulfil the promise of creating a frictionless experience 'in exchange' for the delegation of the photography intention:

> Though Google has gone through great efforts to make getting footage off the Clips camera as painless as possible, it's still a process that involves waiting for the camera to sync with the phone app, sifting through the captured clips, and then moving them over to my phone before I can share them.[60]

The expectation was that delegation would enhance everyday experiences. But this trade-off remained unfulfilled. One critic suggested that the digital images created by the camera were 'lacking in authenticity', as compared to pictures they would have taken themself; the camera was unable to meet the 'like me' expectation that its manufacturers had promised. The technological determinism which largely characterised the commentary by users related to technical, rather that aesthetic or conceptual issues: choices relating to lens, focal length, aperture speed, and other hardware specifications:

> I've been testing the Clips with my two kids for the past couple of weeks, and while I appreciate Google's mission for the product (seriously, anything that gets me to put my phone down more is appreciated), I can't say I'm terribly impressed or happy with the results. **Most of the clips I've been able to capture didn't look better or feel more authentic than what I'm already able to do with my phone or a dedicated camera**.[61] (emphasis by author)

Accordingly, the question regarding whether delegating the prior intention of photography was indeed useful to the users was narrowed down to a more fundamental issue of trust – not in the personal or intimate sense, but in relation to the technical capabilities of the camera: 'Really, the question is how much do you trust the Clips camera and its algorithms to capture the moments you'd otherwise miss?'[62]

The connection between the process of delegating early intent, and the user becoming more rational by virtue of the mediating qualities of Google's self-constructing algorithms, can also be understood from a different perspective: the prior intention, expressed in the personal dimension, to participate in the collective by subscribing to the concept of 'collective intention'. Collective intention, in this context, is not the sum of each individual's intentions; rather, it can be thought of as a 'complex network of intentions and behaviours in which the individual does not necessarily know what the others are doing, but rather believes and acts on the basis of a common purpose and intention for the collective'.[63]

Conclusions

Clips, the autonomous camera designed by Google, could be conceptualised and constructed thanks to the technological affordances created by the Google suite of search, data storage, and software platforms. That aside, the very notion of an autonomous camera is contingent upon a cultural orientation which facilitates delegation to a digital doppelgänger; a mindset which allows for this delegation to be seen as an effective way of negotiating information overload, and ultimately creating a seamless interface linking man and machine. The concept of the Clips camera was predicated on the understanding that over time, users would delegate more and more of their information, actions, and intentions – albeit such as to diminish their agency. Using the autonomous camera reconfigures the notion of prior intention, relegating it from a consciously initiated sequence of actions to a background process. Prior intention, in relation to taking a photograph, was once threaded together by consciously acknowledging the import of a particular event, conceptualising the best means of capturing this as a photographic memory, and initiating the action; Clips proposed to attend to all these, seamlessly presenting the user with the end result of the memory without the need to engage the mediating steps. But, as the user comments indicate, the delegation served to transfer the awareness of prior intention to another dimension. Rather than the actual creation of a digital memory, the user focused on the unsatisfactory technical specifications of the product.

The Clips camera provides a limited case study of the concept of an autonomous camera. One can assert that the concept remains unfulfilled, given that Google, its manufacturer, has withdrawn the product from the market.

That said, its relatively brief appearance on the consumer market, coupled with the small number of users, allowed for a focused analysis of the product from a conceptual perspective. Key observations concerning the Clips camera, relating to intention, mediation, delegation and self-perception, would equally apply to similar products and services provided by Google, given that these technological platforms subscribe to the similar conceptual rationale of the digital doppelgänger.

This article has some limitations. A consideration of the delegation of aspects of the self would ideally engage with aspects of power relations, and the evolving area of 'surveillance capitalism' as related to privacy rights. A more concrete understanding of the technical aspects of creating a digital doppelgänger, such as programming the necessary algorithmic functions to facilitate this, would also be helpful. There is no doubt that products like the Google Clips camera can add value and meaning to everyday life. However, as the interrelationship between technology platforms and digital doppelgängers intensifies apace, future research would ideally be directed towards exploring the impact of these factors on the very notion of human autonomy. We do, after all, have a vested interest in retaining exclusive access to this capacity – or, at least, to be fully aware of the potential consequences of outsourcing it to another agent.

Author

Oshri Bar-Gil is a doctoral candidate in the Psychoanalysis, Culture, and Hermeneutics post-graduate program at Bar-Ilan University, Israel. He is currently writing his dissertation 'Google self: The self-concept at the information revolution'. His research interests are in exploring the ways that people, organisations, and technologies open new horizons for each other to achieve impossible goals.

References

Amadeo, R. 'The Latest Google Shutdowns: Daydream VR, Google Clips', *Ars Technica*, 17 October 2019. https://arstechnica.com/gadgets/2019/10/google-kills-daydream-vr-headset-google-clips-camera/.

Anscombe, G.E.M. *Intention*. Cambridge: Harvard University Press, 1957.

Bar-Gil, O. 'Defining Our Google Self: How Information Technology Mediates Self-Perception For Technology and Self Session at PHTR 2018 Conference' in *Technology and self*. Twente, 2018: https://doi.org/10.13140/rg.2.2.16283.16162 .

Baumeister, R. 'Self-Concept, Self-Esteem, and Identity' in *Personality: Contemporary theory and research*, second edition, Nelson-Hall Series in Psychology. Chicago: Nelson-Hall Publishers, 1999: 339-375.

Belk, R. 'Extended Self in a Digital World', *Journal of Consumer Research*, 40, no. 3, October 2013: 477-500.

Bonnington, C. 'Google's New Smart Camera Isn't Smart Enough', *Slate Magazine* (blog), 27 February 2018: https://slate.com/technology/2018/02/google-clips-smart-camera-isnt-smart-enough-but-its-aims-are-still-worth-considering.html .

Bratman, M. *Intention, plans, and practical reason*. Center for the Study of Language and Information, 1987.

Cakebread, C. 'People Will Take 1.2 Trillion Digital Photos This Year — Thanks to Smartphones', *Business Insider*, 31 August 2017: https://www.businessinsider.com/12-trillion-photos-to-be-taken-in-2017-thanks-to-smartphones-chart-2017-8 .

Castells, M. *The rise of the network society: The information age : Economy, society, and culture*, Volume 1. Chichester-Malden: Wiley-Blackwell, 2010.

Cheney-Lippold, J. *We are data: Algorithms and the making of our digital selves*. New York: NYU Press, 2017.

Clark, A. and Chalmers, D. 'The Extended Mind', *Analysis*, 58, no. 1, 1998: 7-19.

Coeckelbergh, M. *Human being risk: Enhancement, technology and the evaluation of vulnerability transformations*. Philosophy of Engineering and Technology 12. Dordrecht: Springer, 2013.

Erofeeva, M. 'On Multiple Agencies: When Do Things Matter?', *Information, Communication & Society*, 18 January 2019, 1-15.

Feenberg, A. *Heidegger and Marcuse: The catastrophe and redemption of history*. New York: Routledge, 2005.

Floridi, L. *The 4th revolution: How the infosphere is reshaping human reality*. Oxford: Oxford University Press, 2014.

Friese, S. *Qualitative data analysis with ATLAS. Ti*. SAGE, 2014.

Gibbs, G. 'Using Software in Qualitative Analysis' in *The SAGE handbook of qualitative data analysis*. 55 City Road: SAGE Publications, Inc., 2014: 277-294.

Gibson, J. *The ecological approach to visual perception*. Psychology Press, 1979.

Giddens, A. *Modernity and self-identity: Self and society in the late modern age*, first edition. Stanford: Stanford University Press, 1991.

Hongladarom, S. *The online self, vol. 25*. Philosophy of Engineering and Technology. Cham: Springer International Publishing, 2016.

Ihde, D. *Heidegger's technologies: Postphenomenological perspectives*. Perspectives in Continental Philosophy. New York: Fordham University Press, 2010.

_____. *Postphenomenology and technoscience: The Peking university lectures*. SUNY Series in the Philosophy of the Social Sciences. Albany: SUNY Press, 2009.

_____. *Technics and praxis*. D. Reidel Pub. Co., 1979.

Jarrett, K. *Feminism, labour and digital media: The digital housewife*. Routledge, 2015.

Kozinets, R. *Netnography: Doing ethnographic research online*. Reprinted. Los Angeles: Sage, 2011.

_____. *Netnography: Redefined*. SAGE, 2015.

Latour, B. 'On Technical Mediation – Philoshopy, Sociology, Genealogy', *Common Knowledge*, 3, no. 2, 1994: 29-64.

_____. *Reassembling the social: An introduction to actor-network-theory*. Clarendon Lectures in Management Studies. Oxford-New York: Oxford University Press, 2005.

_____. 'Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts' in *Shaping technology/building society: Studies in sociotechnical change*, edited by W. Bijker and J. Law. Cambridge: MIT Press, 1992: 225-258.

Levitt, H., Motulsky, S., Wertz, F., Morrow, S., and Ponterotto, J. 'Recommendations for Designing and Reviewing Qualitative Research in Psychology: Promoting Methodological Integrity', *Qualitative Psychology*, 4, no. 1, February 2017: 2-22.

Liptak, A. 'Google Begins Selling Its Clips Camera', *The Verge* (blog), 27 January 2018: https://www.theverge.com/2018/1/27/16940002/google-clips-ai-camera-on-sale-today-waitlist.

Lovejoy, J. 'The UX of AI – Library', *Google Design Blog* (blog), 25 January 2018: https://design.google/library/ux-ai/ .

Low, C. 'Google Clips Review: A Smart, but Unpredictable Camera', *Engadget*, 27 February 2018: https://www.engadget.com/2018-02-27-google-clips-ai-camera-review.html .

McAfee, A. and Brynjolfsson, E. *Machine, platform, crowd: Harnessing our digital future*. W.W. Norton & Company, 2017.

Menary, R (ed.). *The extended mind*. Life and Mind: Philosophical Issues in Biology and Psychology. Cambridge: MIT Press, 2010.

Mitcham, C. *Thinking through technology: The path between engineering and philosophy*. Chicago: University of Chicago Press, 1994.

Mogg, T. 'Gmail Joins the Billion Users Club', *Digital Trends*, 2 February 2016: http://www.digital-trends.com/web/gmail-joins-the-billion-users-club/ .

Peters, J. 'Google Clips Is Dead', *The Verge*, 16 October 2019: https://www.theverge.com/2019/10/16/20917386/google-clips-dead-discontinued-rip-camera-ai .

Pickering, A. 'Material Culture and the Dance of Agency' in *The Oxford handbook of material culture studies*. 2 September 2010.

Schneier, B. *Data and Goliath: The hidden battles to collect your data and control your world*. W.W. Norton, Incorporated, 2016.

Searle, J. *Intentionality, an essay in the philosophy of mind*. Cambridge [Cambridgeshire]; New York: Cambridge University Press, 1983.

_____. *Making the social world: The structure of human civilization*. Oxford-New York: Oxford University Press, 2010.

Seifert, D. 'Google Clips Review: A Smart Camera That Doesn't Make the Grade', *The Verge*, 27 February 2018: https://www.theverge.com/2018/2/27/17055618/google-clips-smart-camera-review .

Selwyn, N. *What is digital sociology?* Medford: Polity Press, 2019.

Smith, M. 'Intentions: Past, Present, Future', *Philosophical Explorations*, 20, no. sup2, 31 August 2017: 1-12.

Stephens-Davidowitz, S. *Everybody lies: Big data, new data, and what the internet can tell us about who we really are*, first edition. New York: Dey St, 2017.

Vajgel, P. 'Needle in a Haystack: Efficient Storage of Billions of Photos', *Facebook Engineering*, 30 April 2009: https://engineering.fb.com/core-data/needle-in-a-haystack-efficient-storage-of-billions-of-photos/ .

Verbeek, P. 'Cyborg Intentionality: Rethinking the Phenomenology of Human–Technology Relations', *Phenomenology and the Cognitive Sciences*, 7, no. 3, 1 September 2008): 387-395.

_____. *What things do: Philosophical reflections on technology, agency, and design*. University Park: Pennsylvania State University Press, 2005.

Zuboff, S. *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs, 2018.

Notes

[1]   Lovejoy 2018.

[2]   Low 2018.

[3]   Bonnington 2018.

[4]   Liptak 2018; Lovejoy 2018.

[5]   Peters 2019.

[6]   Amadeo 2019.

[7]   Low 2018.

[8]   Cakebread 2017.

[9]   Baumeister 1999, p. 47.

[10] Giddens 1991.

[11] Castells 2010.

[12] Hongladarom 2016.

[13] Searle 2010; Smith 2017.

[14] See, for example, Latour 2005, 1994; Ihde 2009, 2010; Mitcham 1994.

[15] Coeckelbergh 2013, p. 41; Feenberg 2005; Ihde 2009.

[16] Ihde 2009, pp. 87-92.

[17] Ihde 1979; Ihde 2009, pp. 87-92.

[18] Latour 2005; Latour 1992.

[19] Latour 2005; Verbeek 2005.

[20] Latour 1994.

[21] Verbeek 2005.

[22] Zuboff 2018; Schneier 2016.

[23] Mogg 2016.

[24] Kozinetz 2015; Belk 2013.

[25] Kozinets 2011, 2015.

[26] https://www.theverge.com/ , https://www.wired.com/ , https://www.engadget.com/ , https://ars-technica.com/ , https://www.blog.google/ .

[27] Friese 2014; Levitt et al. 2017.

[28] Searle 1983, 2010.

[29] Gibson 1979.

[30] Floridi 2014; McAfee & Brynjolfsson 2017.

[31] Ihde 2009; Verbeek 2008.

[32] Vajgel 2009.

[33] Bar-Gil 2018.

[34] Selwyn 2019; Jarrett 2015.

[35] Ihde 2009; Verbeek 2008.

[36] Latour 1992, 1994, 2005.

[37] McAfee & Brynjolfsson 2017.

[38]   Gibson 1979.

[39]   Latour 1992.

[40]   Clark & Chalmers 1998.

[41]   Clark & Chalmers 1998; Menary 2010.

[42]   Anscombe 1957; Searle 2010; Smith 2017.

[43]   Searle 1983, p. 1.

[44]   Bratman 1987.

[45]   Smith 2017, p. 1.

[46]   Searle 2010.

[47]   Bratman 1987; Searle 1983, 2010; Smith 2017.

[48]   Seifert 2018.

[49]   Gibson 1979.

[50]   Seifert 2018.

[51]   Bratman 1987.

[52]   Latour 2005.

[53]   Erofeeva 2019, p. 4; Latour 1992.

[54]   Pickering 2010.

[55]   Searle 2010, p. 133.

[56]   Bratman 1987.

[57]   Seifert 2018.

[58]   Cheney-Lippold 2018.

[59]   Seifert 2018.

[60]   Ibid.

[61]   Ibid.

[62]   Ibid.

[63]   Searle 2010, p. 60.