‘DIY AI’? Practising kit assembly, locating critical inquiry[1]

Philippe Sormani

University of Lausanne / CEMS-EHESS, Paris

Abstract

This paper presents a reflexive ethnography of ‘DIY AI’ underway. Part 1 examines a promotional video of Google’s ‘AIY Vision Kit’, its ‘do-it-yourself intelligent camera’, running on a Raspberry Pi computer and fitting into an 4.7×7.5×7.6 cm cardboard box. Part 2 of the paper, in turn, reports on our initial effort at kit assembly with the help of the user manual. In particular, I shall home in on our ‘turn it on’ attempt, as a first ‘step condition’ to operate the assembled kit ‘intelligently’—that is, for ‘experiment[ing] with image recognition using neural networks’ (Google 2018). The reflexive ethnography pursues two aims. First, it shall make explicit (some of) the ‘vulgar enabling practices’ (Button and Sharrock 1995) of the probed ‘intelligent camera’. Second, the ethnography will revisit the interplay between ‘technical work and critical inquiry’ (Lynch 1982) by locating how, when, and why the former invited the latter *in situ*. Recent reflection on critical inquiry in and across STS (e.g., Mirowski 2020), algorithm studies (e.g., Mackenzie 2017), and social and cultural studies more broadly (e.g., Tsilipakos 2018), will be practically indexed and recast accordingly. So will critical inquiry in matters of ‘DIY (AI)’ more specifically.

[1] Prior versions of this paper have been presented on different occasions, including the IIEMCA 2019 conference in Mannheim, Germany, an international workshop on digital governance at the EC Joint Research Centre in Ispra, Italy, later that same year, and a recent online data session on ‘AI in Interaction’. I wish to thank all participants for their helpful remarks and critical observations, as well as Alain Ravenel, Daniele Belitrandi, and Phillip D. Brooker for having been variably involved in the fieldwork leading up to this paper (as the first of a three-part working paper series, see Sormani, forthcoming a, b). The finalization of the paper benefitted from the additional observations by the editors of this special issue, two anonymous reviewers, Hannah Pelikan, and Oskar Lindwall, and was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 262513311—SFB 1187. Remaining mistakes are mine.

INTRODUCTION: UBIQUITOUS AI, REFLEXIVE ETHNOGRAPHY, AND CRITICAL INQUIRY

> 'Is it possible for a machine to think?' […] [T]he trouble which is expressed in this question is not really that we don't yet know a machine which could do the job. […] The trouble is rather that the sentence, 'A machine thinks (perceives, wishes)' seems somehow nonsensical. It is as though we had asked 'Has the number 3 a colour?'' (Wittgenstein 1960:47).

For Wittgenstein, the 'thinking machine' was a conceptual absurdity, and so it has remained for many of his most informed commentators (e.g., Button et al. 1995; Shanker 1995). Regardless of the conceptual absurdity though, the research program of 'artificial intelligence' (AI) has given rise to numerous computer systems, applications, and devices which today are promoted and/or used *as if* they were 'smart', 'skilled', or 'intelligent' somehow. AI, in short, has become ubiquitous—a virtually ubiquitous engineering success. Internet search engines, such as 'Google', and their automated recommendation systems (on the basis of previous search history) provide a case in point. So do many 'smartphone' applications (e.g., Techgig 2020). Over the last ten years, the multiplication of 'intelligent systems' moreover has given rise to innumerable reports, books, articles, and opinion pieces charting the phenomenon, be it from a technological or economic, political or ethical, historical or sociological perspective. Often these perspectives and those genres are difficult to disentangle, and inter- or transdisciplinary approaches are called for to probe their intricacies (e.g., for the purpose of 'governance', Bloom 2020).[2]

This paper recasts the 'curious incongruity' (Garfinkel 2002) between conceptual absurdity (or conundrum) and engineering success in contemporary AI as a topic of empirical inquiry. For the purpose, the paper presents a reflexive ethnography of 'DIY AI' underway, taking a closer look at how a particular 'AI' system was promoted and could be practiced as a 'DIY' endeavour. The system in question is Google's 'AIY Vision Kit' (2018), its 'Do-it-yourself Intelligent Camera' (*ibid.*), running on a Raspberry Pi computer and fitting into an 4.7×7.5×7.6 cm cardboard box (see Fig. 1).

[2] For a detailed discussion of 'machine intelligence' as a conceptual conundrum, see Shanker (1988). The conundrum, if not absurdity, is presented as an intricate mix of conceptual assumptions and disentangled in terms of its main components, including mathematical formalism, neural network modeling, lingering behaviorism, algorithmic reductionism (from practical rules to causal mechanisms), cybernetic feedback looping, and a stipulated 'learning continuum' (across human- and non-human 'agents', from their 'simplest' to most 'complex' forms). The question of how the mix plays out over the course of AI history is beyond the scope of this paper. Cardon et al. (2018) allude to the recent revival of 'machine learning' in AI, and the 'fabrication of neural networks' in particular, as 'a kind of black magic' (*une sorte de magie noire*) (p. 27). The allusion suggests that the conceptual muddle, if not engineering challenge, is still with us (e.g., Lake et al. 2017).

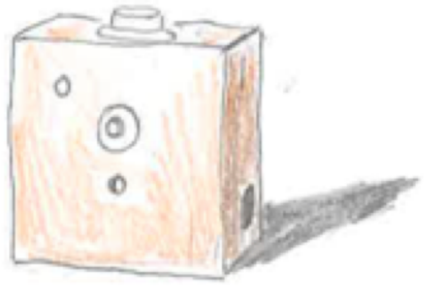


Fig. 1: Google's 'AIY Vision Kit' as staged in the user manual (drawing by the author)

As part of Google's 'AIY: Do-it-yourself Artificial Intelligence' (2020a) product line, this 'intelligent camera' was advertised in 'DIY' terms: 'With our maker kits, build intelligent systems that see, speak, and understand. Then start tinkering. Take things apart, make things better. See what problems you can solve' (Google 2020a). The ensuing ethnography, in particular, will home in on Google's 'AIY Vision Kit' by addressing the following questions: how was the 'Vision Kit' promoted as a cogent 'AI' device? How could it be practically assembled, if not enabled, via a dedicated 'DIY' effort? And how would kit promotion (in Google's 'AIY' terms) and kit assembly (via our 'DIY' effort) relate? The ethnography will take on a 'reflexive' character in an ethnomethodological sense, insofar as it will inquire into the *ordinary availability* of the 'Vision Kit' to the analyst, whilst explicating some of the mundane procedures (i.e., 'ethno-methods') which provided for that availability in a promotional video to begin with, and then through actual kit assembly.[3]

How will 'critical inquiry' fit into the picture? In the process of conducting the ethnography, I noticed a familiar ambivalence: the 'DIY' theme turned out to be both a clarifying and obfuscating resource in dealing with 'AI' as a 'computational artifact' (Schmidt and Bansler 2016). On the one hand, the promotional video of the device in 'DIY' terms obfuscated the practical challenge as it staged the 'intelligent camera' as an easily to be assembled kit, requiring no professional or particular computational skills in the process (an ordinary 'workshop setting' would do). On the other hand, the multiplication of practical problems encountered in kit assembly made it clear, at least to me as the involved ethnographer, that local tinkering activity ('DIY') remained a far cry from smooth product customization ('AIY'), not to mention a practical engineering solution to the perennial conceptual conundrum ('AI'). The noticed ambivalence proved 'familiar' in two respects: not only did my trials and tribulations in 'DIY AI' resonate with

[3] As R. Turner put it half a century ago, 'the task of the [ethnomethodological] analyst in analyzing naturally occurring scenes is not to deny his [society-member's] competence in making sense of activities, but to *explicate* it' (Turner 1970:197; emphasis in original). For a recent introduction to ethnomethodological analysis in this vein, see Sormani (2019). In that introduction, I argued for (ethno-)methodological 'triangulation' as reflexive heuristics, where one approach (e.g., technical self-instruction) may improve upon another (e.g., video analysis). This paper takes its cue from said triangulation, whilst working at its possible extension (i.e., by adding initial observations made in the course of (re-)drawing screenshots).

other 'digital amateurs' (in August 2020, I noted '559 [amateur] issues' with the 'Vision Kit' on *github*, a software self-help website), but those trials and tribulations also resonated with the ambivalent character of 'DIY' initiatives, digital or analogue, as discussed in the academic literature, approaching them as an ambivalent matter of critically engaged 'adversarial design' and/or tacitly consensual public relations (e.g., DiSalvo 2012; Ratto and Boler 2014). The paper will conclude on that discussion, after having revisited the interplay between 'technical work and critical inquiry' (Lynch 1982) by locating how, when, and why the former invited the latter *in situ*.[4]

PART 1—PROMOTIONAL VIDEO: 'THIS YEAR WE ARE RELEASING THE VISION KIT' (2017)

In summer 2018, I landed on Google's 'AIY: Do-it-yourself Artificial Intelligence' projects website (Google 2020a) and, upon an online order and UPS home delivery totaling circa 110$, received the purchased item two weeks later in my mailbox: the 'AIY Vision Kit' (Google 2020b). This 'Vision Kit' was released in late 2017, as the second 'AIY' project, after a first 'Voice Kit' (Google 2020c) released in May that same year (cf. Rutledge 2017). No further kits have been launched since. That being said, and with the 'Vision Kit' in hand, why should I bother examining its promotional video? Simply put, I first came across the 'Vision Kit' through that video, as it was to be found via Google's 'AIY' projects website. Not only was the kit's 'ordinary availability' video-based to begin with, but its promotional video persuaded me to get the kit. How could that be? In answer to this question, Part I of this paper draws upon a detailed transcription of the promotional video of Google's 'AIY Vision Kit' (2020b). In particular, the ensuing analysis shall probe its audiovisual montage in making explicit how the 'Vision Kit' was promoted through this selfsame montage, for the kit's technical analogue of 'intelligent seeing'—that is, 'image recognition'—to prove intelligible and interesting to a broader audience of 'digital amateurs'.[5]

---

[4] Lynch's study, his first publication in *Social Studies of Science*, highlighted how technical work afforded laboratory scientists with empirical opportunities for critical inquiry (p. 511). Insofar as they were pursuing these local opportunities, scientists didn't appear to stand in need of critical enlightenment by a disengaged scholar. Their technical work already contained recognizable courses of critical inquiry. Of late, however, critical inquiry has reemerged as a topic of academic interest, as recent lines of reflection in and across STS (e.g., Mirowski 2020), algorithm studies (e.g., Mackenzie 2017), and social and cultural studies more broadly (e.g., Tsilipakos 2018) suggest. This paper takes a first stab at (re-)locating these lines of reflection, and critical inquiry in matters of 'DIY (AI)' more specifically.

[5] In the introduction, I emphasized the 'non-intuitive' character of AI as a research program, insofar as it proceeds from an (arguable) category-mistake implying the attribution of mental predicates to machines (for a recent discussion, see Brooker et al. 2019). Wittgenstein's 'trouble' might or might not be widely shared these days, it remains interesting for methodological purposes as it raises the question of how 'machine intelligence' is presented as an intelligible, if not interesting, endeavor to a broader audience—the question that the ensuing video analysis will address. The expression 'digital amateurs' stands as a shorthand for the manifestly intended recipient of the promotional video (i.e., persons interested in assembling the 'Vision Kit' for themselves). As T. Dant explains, 'DIY usually refers to work on the shaping and structure of the domestic environment […]', including "niche' DIY cultures such as those around electronics and digital equipment with manuals and magazines as well as support groups' (2019:285).

1.1. Delivering the news

For the kit to be promoted, its promotional video needs to be understood. One way to make possible a broad, amateur understanding is to draw upon, factor in, and articulate the same kind of 'cultural knowledge' that society-members regularly draw upon in everyday interaction, or at least can be presumed to do (cf. Jayyusi 1988:272). The promotional video of present interest draws upon two sets of such cultural knowledge: first, its audiovisual articulation in so many 'say-shows', an ordinary resource from classroom instruction (e.g., Burns 2012); second, mundane patterns of conversational interaction, such as 'news delivery sequence[s]' (Maynard 1997) or 'initiation-response-evaluation' sequences (also to be found in classroom interaction; cf. Mehan 1979). The video announcement of the 'Vision Kit' (see Excerpt 1) draws upon all of these procedural resources, and more, whilst articulating them in a particular way (also see Sormani 2020, upon which the following section draws in part).[6]

Shot	*Voice*	*Audio*	*Video*
1	Maker Narrator Maker	*crac, crac* ((puts box on table)) ***#1*** **this year, we are releasing the,** ((opens box))	 ***#1*** (('Vision Kit' box, shown close-up))
2	Narr Mak	**vision kit.** ((installs camera)) *zac* ***#2***	 ***#2*** (('Vision Kit', shown close-up))

[6] The present video analysis uses the following transcription conventions:

=	latching, no discernible interval between adjacent utterances, or activities
she-	cut-off
he: ::	elongation (successive colons indicating increased elongation)
he	emphasis, louder voice
.	falling intonation
,	'continuing' intonation
(go ahead)	uncertain hearing or seeing
((does))	description, comment
#1	indication of screenshot placement in the transcribed activity
zac	tinkering noise

3	Mak Narr	((shows mini-computer)) ***#3*** **it does not need=** **>an internet connection,<** **it does not need the cloud.** **y=**	 ***#3*** ((Maker shows mini-computer))
4	Narr Mak	**=ou (.) program it, wi:th** ((closes 'Vision Kit' cardboard box)) *chic, chic* ***#4***	 ***#4*** ((Maker closes 'Vision Kit'))
5	Narr Mak Narr	**the neural network model.** ((silently turns white 'Vision Kit' button towards camera)) ***#5*** **and,**	 ***#5*** ((close-up on white button on 'Vision Kit'))
6	Narr Narr Narr	**the kit then,** **gives you back,** ((gestures toward himself)) ***#6*** **eh=>data,<=on what**	 ***#6*** ((close-up shot of narrator, appearing as maker now))
7	Narr Mak	**it sees.** ((silent 'slide in' of joy detector label from left)) (((silently turns on 'Vision Kit'), which lightens up in blue)) ***#7***	 ***#7*** ((shot on top button of 'Vision Kit', now lightened up in blue))

8	Narr Narr	**so=>if i give it a<=face,** **#8** **the light will come on,**	 **#8** ((close-up shot of narrator/maker looking at 'Vision Kit', still lightened up in blue))
9	Narr Narr	**(.) and,** **#9** **as i get happier,** *bliip*	 **#9** ((close-up of 'Vision Kit', still lightened up in blue))
10	Narr (Mak)	((changes facial expression into 'happy face', silently stares at 'Vision Kit')) ***#10***	 ***#10*** ((shot of 'smiling' narrator/maker))
11	 Narr Narr	*>blip, blap, blip<* **the light will change,** ***#11*** **from blue to yellow.** ((silent 'slide out' of joy detector label, towards left))	 ***#11*** ((reshown 'Vision Kit', now close-up))

Excerpt 1

Notice first how the announcement of the kit is articulated: not only does it take the form of an indicative sentence, stating a putative fact ('this year, we are releasing the…', shot 1), but its projected object—'… [the] vision kit' (shot 2)—is shown-as-it-is-stated—that is, it is shown to be assembled at the very moment it is mentioned, with the camera objective being mounted on its cardboard box ('*zac*', shot 2, ***#2***). Moreover, the successive shots, organized in terms of as many 'say-shows', articulate a particular 'news delivery sequence' (Maynard 1997). From ordinary conversation, the sequence borrows its constitutive parts: an announcement, followed by its response, an elaboration, and a final

assessment (*ibid.*, p. 97). Yet the parts are not distributed as reciprocating turns at talk, as in conversation, but across the successive 'say-shows' that compose the audiovisual montage. Accordingly, the 'announcement' is first stated ('this year we are releasing the vision kit', shots 1 & 2), for the viewer then to be able to arrive at a 'response', a potentially positive one, as (s)he is being shown the assembled kit in silence, for a second, after its final piece—the camera objective—has hearably and visibly been added ('*zac*', shot 2, ***#2***). The subsequent 'elaboration', in turn, spells out the kit's operating conditions (shots 3–5) and regular operation (shots 6 and 7).[7]

The operating conditions, as they are verbally elaborated upon, are shown to already have been established: a small piece of computer equipment is shown (shot 3, ***#3***) at the moment a series of unnecessary operating conditions are listed (shot 3), implying that the former equipment by contrast already secures the necessary ones. This contrastive implication, then, is swiftly spelled out in terms of computer programming ('you program it with the neural network model', shots 3–5) which in turn is suggested to have already been successfully implemented (as the closing of the kit's cardboard box is shown, followed by a close-up on the whitish, transparent button on the kit's top, shots 4 and 5, ***#4*** and ***#5***). Conversely, the regular operation of the kit is not shown right away, but simply stated and shown to be stated (as the utterance 'the kit then gives you back data on what it sees' is shown to be made by the narrator, incidentally appearing as the involved 'maker' too, shots 6 and 7, ***#6***). This temporary withholding of kit demonstration, whilst tying the verbal elaboration to the workshop setting, begs the question of the 'final assessment' by the viewer of the conveyed news item: '[this year's] vision kit' (shots 1 and 2, ***#2***). The sequel of the video seems to be both anticipating this expectancy and facilitating its fulfilment.[8]

The expectancy of a 'final assessment', to bring to completion the 'news delivery sequence' oriented to as a normative matter, appears to be implied in the early display of the operating kit (labelled as a 'joy detector' and shown with its lightened-up button, shot 7, ***#7***). Its display appears as 'early', insofar as it is made before the verbal explanation of the kit's operation, as shown to be uttered in the workshop, comes to a close (shots 6 and 7). At the same time, the verbal explanation's ending on the utterance's subject and predicate ('… it sees', shot 7), as it is paired with the kit's manifest operation (shot 7, ***#7***), facilitates the viewer's expected 'final assessment' by providing and specifying visual grounds for it (as the operating kit is shown, and labelled as the 'joy detector', shot 7, ***#7***, video). This assessment is further facilitated by the ensuing demonstration, insofar as the

---

[7] Technical work, to reiterate Lynch's (1982) observation, offers its practitioners a locally available set of empirical affordances for critical inquiry (p. 511). In the present case, the detailed transcription of the audiovisual montage makes possible critical inquiry of how its 'say-show' parts relate to each other, both in terms of their intelligible articulation and noticeable inconsistencies. In what follows, I shall list some of these possible inconsistencies, as the footnoted subtext to the video analysis of that manifest articulation.

[8] Noticeable inconsistency #1: 'the kit then gives you back data on what it sees' (shots 6 and 7)—does it? This question, raised by a colleague during our online data session, challenges the video for not 'showing' what it states (or, at least, not doing so in an obvious way). At best, the video shot (*#7*) can be seen to show the operating kit (given its lightened-up button), yet without explaining what 'data feedback' means, might or will mean in this instance (Oskar Lindwall, personal remark).

latter persuasively trades upon another familiar conversational sequence, an 'initiation-response-evaluation' (IRE) sequence (shots 8–11).

The persuasive character of kit demonstration lies in the sequence's partial formulation—that is, the 'initiation' and 'response' are verbally formulated ('so if I give it a face, the light will come on', shot 8), whilst the 'evaluation' is implied on visual grounds (as the lightened-up and 'joy detector'-labelled kit is shown, shot 8, ***#8***) and then suggested to be left to the viewer on those grounds (as the lightened-up 'joy detector' is zoomed in on and, thereby, shown to the viewer exclusively, shot 9, ***#9***). To accomplish kit demonstration, the pattern is repeated: not only a partial formulation of the IRE sequence is repeated (i.e., of the I and R parts first, as in 'and as I get happier, the light will change, from blue to yellow', shots 9–11), but also the pending 'evaluation' (E) is suggested to be left to the viewer again (as the lightened-up 'joy detector', now 'yellow', is zoomed in on then and, thereby, shown to the viewer exclusively again, shot 11, ***#11***). A positive 'final assessment' of his (or hers) of the news item under consideration is thus implied.[9]

1.2. Managing contingencies

In the 'perceptual world of activity and interaction', to use Jayyusi's phrase (1988:272), conversationalists 'shape each component [of a news delivery sequence] according to a *myriad of contingences*', as Maynard points out (1997:98; emphasis added). Retrospectively, a 'central contingency in the development of [...] news […] is a recipient's prior knowledge of the occurrence to be reported' (Maynard 1997:95). Prospectively, a central contingency of a news item, in and upon the course of its development, is its projected ensemble of practical implications for the recipient. In conversation, the retrospective contingency is typically dealt with by way of a 'preannouncement' (Maynard 1997:95–96). In the case of the promotional video, the preannouncement of the 'Vision Kit' (prior to the examined episode) happens to be built out of several components, across several 'say-shows', including the incidental display of the kit's packaging, the staging of the 'AIY projects team' and '[teenage] makers' being involved in kit development, and its 'intelligent' predecessor, the 'Voice Kit' (see Google 2020a, Vision Kit promotional video, seconds 1–30). Taken together, these components articulate a first set of 'felicity conditions' for the subsequent news delivery sequence to appear broadly intelligible, if not technically persuasive (as analysed in our 'Vision Kit' case so far). Prospectively, in turn, a central contingency is projected practicality. Accordingly, the promotional video of the 'Vision Kit' has its first demonstration of the kit's 'joy detector' function backed up with

[9] Noticeable inconsistency #2: '[…] as I get happier, the light will change from blue to yellow' (shots 9-11)—does it? Upon closer inspection, the audiovisual montage can be seen, not to show said 'color change', but to juxtapose the kit button lightened-up in 'blue' (shots 9 & 10) and, one shot later, in 'yellow' (shot 11). The color change appears as a montage effect, rather than or regardless of its material production. For a related analysis, see Ball and Smith (2019).

two further demonstrations: a demonstration of its 'object classifier' and 'face tracker' functions, respectively (Google 2020a, *ibid.*, from second 58 onwards).[10]

PART 2—KIT ASSEMBLY: 'TURN IT ON', OR HOW TO MAKE THE (VISION) KIT WORK

To have a technical device 'shown to work' in a promotional video is one thing; to have a technical device 'made to work' so that it can be shown is quite another, be it for promotional or other purposes. The preceding video analysis reminded us of that 'vernacularly available distinction' (Garfinkel 2002:198). The promotional video of Google's 'Vision Kit' offered an apt opportunity for shot-by-shot analysis of its audiovisual montage of technical performance. In the process, we noticed how and why the video demonstration did not require the performance's 'material production', but did stage it as a 'montage effect', an interactionally contingent and persuasively designed one at that. At the same time, we noticed a first set of conversational contingencies (e.g., a multi-component 'preannouncement') taken into account, in and for the promotional video. Part II of this paper, in turn, will ask 'what more can be learned' from trying to assemble the 'Vision Kit' in actual practice, in addition to the video analysis of its persuasive montage regardless of such practice. To address this question, I will proceed from another vernacularly available distinction, the distinction between a 'textual instruction' (e.g., in a user manual) and its 'practical engagement' (via manual use).[11]

2.1. Co-assembling the kit

Google's 'AIY' projects website (2020a) and its 'Vision Kit' promotional video (*ibid.*) both pitched 'DIY AI' as a collaborative endeavour, involving a 'projects team' and a 'maker community' alike. The promoted tinkering experience also appeared as part of an 'hands-on' amateur activity in and yet beyond the digital domain, rather than a professional pursuit restricted to it (e.g., similarly to bike repair via Youtube, Dant 2019).

[10] Noticeable inconsistency #3: 'so if I give it a face, the light will come on' (shot 8)—does it? Again, and upon closer inspection, the audiovisual montage does not show what it states (indeed, the kit's 'light' is already 'on', as the narrator is shown to 'give it a face', shot 8, *#8*). Taken together, this inconsistency and the previous one (see note 9) specify the difference between the persuasive use of a conversational structure (e.g., an IRE sequence) and the material production of an 'intelligent machine' (i.e., the IRE sequence perhaps offers an interesting, but insufficient condition for this latter purpose). For a recent attempt to fuse the two, see Moore and Arar (2019).

[11] In other words, my trials and tribulations as a 'digital amateur' will offer us another tack to revisit Lynch's (1982) plea for examining the interplay between 'technical work and critical inquiry'—that is, how, when, and why the former invited the latter *in situ*. Accordingly, the list of noticeable inconsistencies will not only be prolonged, but the status of its proliferating items will eventually be questioned, too. When (re-)drawing the screenshots for this paper, I noticed a series of incongruities suggesting in what direction. For one thing, I noticed the stark contrast on the 'Vision Kit''s lens between black and white (see shot *#2*), a contrast which seems to reflect a professional photo studio more so than an amateur workshop. For another, I noticed the numerous hammers on the workshop wall behind the smiling narrator cum kit demonstrator (see shot *#6*). For kit assembly, these hammers *prima facie* didn't seem useful, if only for the dimensions and materials of the 'Vision Kit'. Incidentally, the hammers suggest another move, if not artistic performance: kit smashing!

Accordingly, and upon receiving the 'Vision Kit' at home, I located promising partners for kit assembly among friends and acquaintances to start with, including an arts and crafts teacher met at a friend's cocktail party (never mind how for now, but see Garfinkel and Wieder 1992:182–183). His involvement should facilitate kit co-assembly, or so I thought. Our first session at least did not suggest problems ahead, but a swift and successful opening (see Vignette 1).

> In autumn 2018, I teamed up with Rod Alexis, an arts and crafts teacher at a primary school in the Montreux area, Switzerland. We teamed up to assemble Google's 'AIY Vision Kit', its 'do-it-yourself [artificially] intelligent camera, [allowing you to] experiment with image recognition using neural networks,' as its *User Manual* explains.
>
> After a first meeting at university (I got to know Rod at a friend's cocktail party), we meet at the local school where Rod teaches, on a Wednesday afternoon, as its pupils' half-day off. I bring along the 'AIY Vision Kit', and it takes us about 50 minutes to assemble it in 69 working steps. The *User Manual* congratulates us at step 70: 'Congrats, you've just assembled the Vision Kit hardware! Now you're ready to turn it on.'

Fig. 1.1

Fig. 1.3

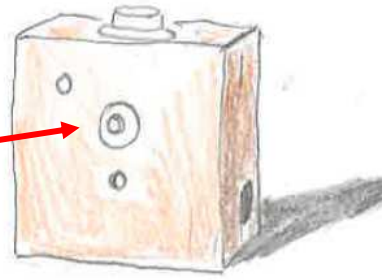


Fig. 1.2

> Apart from the background (reminiscent of the workshop staged in the promotional video), the picture of the device assembled *in situ* (Fig. 1.1) looks the same as the one shown in the user manual (Fig. 1.2). Hence the double-arrow (Fig. 1.3).

Vignette 1

The opening session, as the above vignette suggests, allowed us to swiftly assemble both the 'Vision Kit' and the workshop setting suitable to its assembly. The setting (shown in Fig. 1.1) appears as a cameo version of the workshop staged in the promotional video. At least, it appears to be patterned in terms of 'cultural particulars' partly similar to those made available by that video. On the one hand, 'DIY' tinkering appears as a 'hands-on' endeavour, too: not only the open packaging of the kit is shown, but also our manual grasp of it suggests practical engagement and tactile kit assembly (as in the video,

Excerpt 1, shots 1–4). So do the taken-off glasses (in the middle of Fig. 1.1). On the other hand, the assembly team has changed: I'm described to be engaging in kit co-assembly, not with a likeminded 'MIT alumnus' or 'teenage makers' (cf. Google 2020a, b), but with a Montreux-based 'arts and crafts teacher' (Vignette 1). In so doing, we can be seen to be expanding 'maker culture' (given the school setting in a foreign place, at least with respect to the US), although the future incumbents of that culture are not physically present (it is a 'Wednesday afternoon, […] pupils' half-day off', *ibid.*).[12]

In any case, the 'Vision Kit hardware' appears to have been successfully co-assembled during the reported Montreux-session, not to have it 'shown off' for promotional purposes, individually or collaboratively, but to get 'ready to turn it on' (as the user manual suggests at 'step 70', Fig. 1.2). Taken together, the recognizably re-enacted workshop setting and the manifestly assembled kit led me to harbour a 'constancy hypothesis' across time and space. *Across space*: the locally assembled 'hardware' would be the same, and it would be working in the same way, as the one shown in the user manual (cf. Figs. 1.1, 1.2 and 1.3). *Across time*: this identity assumption, the assumption of the 'working kit', would continue to hold, potentially indefinitely (e.g., a year later).[13]

2.2. Reconsulting the manual

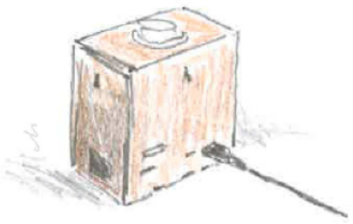
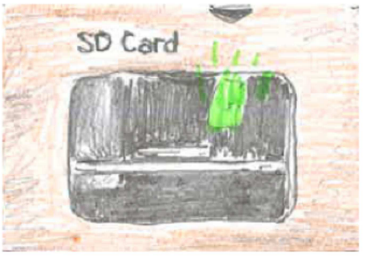
On the basis of my identity assumption, I reconsulted the user manual the following summer and, whilst doing so, re-instantiated that very assumption. Retrospectively, I assumed that manual consultation so far had allowed us to correctly assemble the kit, the 'real thing', not simply a staged version of it. Prospectively, I got back to both this 'real thing' and the user manual, whilst attempting to use the latter to bring to life the former (i.e., to make the 'Vision Kit' work). The manual, in other words, was not solely assumed to have provided us with a 'context of looking' (Baccus 1984:38), but the manual's following in terms of that selfsame context, as the 'ideal account of proper procedure' (*ibid.*), was assumed to have been productive, productively kit constitutive (literally 'step by step'). No readily recognizable 'dumb move' (*ibid.*) had been made in the opening session, and no such move was expected for the 'steps ahead' (as projected via the manual and its manifest use). That, at least, was my stance culled from the opening session, and tentatively played out to meet '*kit operating condition no. 1*' (i.e., to 'turn it on'), back at the office in June 2019 (see Vignette 2 below).

---

[12] Noticeable inconsistency #4: 'Google's 'AIY' projects website (2020a) and its 'Vision Kit' promotional video (*ibid.*) both pitch 'DIY AI' as a collaborative endeavor'—do they? As I engaged in tentative kit co-assembly, I noticed that the website and video variously alluded to a 'collaborative endeavor' (e.g., by mentioning the 'AIY projects team' or staging '[teenage] makers' observing a running kit), whilst showing kit assembly as an individual accomplishment (e.g., Excerpt 1). The user manual seems to imply the latter, too (cf. Google 2018).

[13] Why? This was the question by a first reviewer. How? This was a question by a second reviewer. Having managed to assemble the kit in a 'workshop setting' similar to that shown in the promotional video, I manifestly assumed the kit to be operational as suggested by that same video. This manifest assumption is similar to taking for granted the 'ordinary circumstances' of a performative utterance, which apparently I did on the basis of the promotional video (for further elaboration of this point, see Sormani 2020).

'Let's turn it on!'—this was my confident thought, back at the office, as I started toying around with the assembled 'AIY Vision Kit' in June 2019. 'Let's have this Kit run'—from 'step 71' onwards, at least towards and including 'step 72' and 'step 73' of the *User Manual*. For this purpose, each step of the manual provides an instruction articulated into a readily followable 'show-say' pair—first, the step 'shows' its own realization, then it 'says' how to accomplish it. Here are the steps 71, 72 and 73 as 'shown-said' in the manual:

Try it out **TURN IT ON**	
	**71. Plug your Vision Kit into a power supply** Plug your Vision Kit into a wall power supply through the port labelled Power on your device. See Meet your kit for power supply options. Do not plug your Vision Kit into a computer for power.
	**72. Let it boot up** To confirm that it's connected to power, look into the hole in the cardboard labelled **SD Card**. You'll see a green LED light flashing on the Raspberry Pi board. Be patient while it boots up; the first boot takes a few minutes. You'll know it's booted when you hear a short tune. The software needs this time to install and configure settings. In the future, it'll start faster. When you decide to put away your kit, follow the steps to safely shut it down.
USE THE JOY DETECTOR	
	**73. Try out the Joy Detector** Point the Vision Kit toward someone's face (or your own) to try out the Joy Detector demo. When the camera detects a face, the button illuminates. – Ask them to smile – Then ask them to smile REALLY BIG – Then ask them to make a frowny face The Joy Detector uses machine learning to detect if a person is smiling or frowning, and how much they are doing so. A smile turns the button to yellow, and a frown turns it blue. […]

> The successive 'show-says' suggest 'plug-and-play' availability, conveyed via a noticeably gendered, ageist, and individualist framing. Yet this suggestion is only the textual instruction and its visual display, the latter preceding the former (at least in terms of culturally prevalent reading rules, from left to right, and titles being set aside).
>
> *Vignette 2*

So I reconsulted the manual. Vignette 2 above elaborates both on that reconsulting episode (in the past tense) and the reconsulted manual (in the present tense). In and as those reported material specifics, the vignette implies the mundane distinction between the 'textual instruction' and its 'practical enactment'. Let us pause on the former, presently 'available in disengageable text' (Garfinkel 2002:199), before turning to its practical enactment *in situ*.

Although not stated as such, '*kit operating condition no. 1*' is suggested to be reachable upon the joint completion of manual steps 71 and 72. Not only step 72, and the 'green LED light flashing' in particular, mark the accountable success of step 71, but the realization of this operating condition is suggested to be 'relatively easy': for one thing, the two steps are introduced subsequently to, and conditionally on, the cogent realization of seventy prior steps (!), as assumed by the congratulatory manual and the confident user (see Vignette 1), as well as the assembled 'hardware' shown (as the turned-around, yet identical 'working kit' again; at step 71, Vignette 2). For another, the implied user is supposed to have moved into a position of competent kit use, pending a first '[software] boot up' (more of which below) and equipped with additional instructions to 'safely shut it down' (at step 72, *ibid.*), all of which prior to the 'use the joy detector' section (which in turn assumes the 'turn it on' operating condition to have been met, at step 73, *ibid.*). So much for manual re-consultation, in and as its 'disengageable text'.[14]

2.3. Testing the kit

For the 'Vision Kit' to work, its prospective user has to be able to 'turn it on'. In the promotional video, this first kit operating condition is shown to hold after a few seconds, as its protagonist is shown to press the button with his right hand on top of the 'Voice Kit', which then is shown to run as expected (see Google 2020a, Vision Kit promotional video, seconds 1–30). For the 'Vision Kit', the same condition is suggested and shown to hold in that same video only a few seconds later, as its button in turn lightens up (Except

[14] Noticeable inconsistency #5: the user manual's 'show-say' instruction seems to be inconsistent with the promotional video's 'say-show' montage—does it? The short answer is 'no' (promotion is not instruction). A longer answer might be 'yes', and goes as follows. The *user manual's consistently used 'show-say' instructional format* suggests 'relatively easy' practical realization of each step required to have 'the kit work'. The priority put on 'showing' each reached step, rather than on 'saying' its realization in detail, manifests a 'minimization rule', as the image shown can be compared as directly as possible with the kit assembled *in situ* (i.e., by minimizing the user's reading of the manual's textual part). Conversely, the *promotional video's persuasive 'say-show' montage* suggests a 'maximization rule' to be at play, rule according to which 'whatever is said' is to be aggrandized by 'what is shown' (i.e., to be maximizing the viewer's expectancy that, indeed, what is said is shown). Hence, an ironic gap between aggrandizing promotion and agonizing practice may open up.

1, shot 7, **#7**). The manifest display of facile kit operation, be it in the promotional video or in the user manual, begs the question of the practical realization of the displayed operation—its 'practical enactment' or, put otherwise, its 'instructed action' (Garfinkel 2002:197). How might I, as a prospective user, make the kit work and, more specifically, make it intelligibly and consistently work, in and as its 'facile operation'? Consider Vignette 3 in answer to this question.

> Wednesday afternoon, June 2019, at the office. It was a couple of months ago that I last touched the 'AIY Vision Kit,' so I pause to parse: what do I actually need, and what can I leave aside, to have the 'Kit' up and running? What are the *relevant* 'experimental conditions', if, say, I'm going to try out the 'Joy Detector'? Here's my *in situ* list of 'no needs' to begin with:
>
> - There's no need for a 'transformer'—the US power supply cable, which was delivered with the 'Kit' in my mailbox, covers a 110–240 Volt range, so it should be fine for my European use, too (at 220 Volt).
> - There's no need for a 'workshop' either—the 'AIY Vision Kit,' including its camera, run on a 'Raspberry Pi, a low cost, credit-card sized computer'—both, camera and computer, being packed into the cardboard box. No need for a room-filling mainframe computer as in the 1950s!
> - There's no need for a 'collaborator' at this stage—indeed, me and the manual should be just fine, to have the 'Kit' up and running, and its 'Joy Detector' work. No need to crack (e.g., Python) code yet.
>
> Out of this '*in situ* list,' while pondering its items, I cull the decision to go ahead with a first test, a 'first experiment.' Therefore, I assemble the 'Vision Kit,' the 'US power supply cable,' a 'EU power adapter' (plugs are not the same) into a 'testable configuration.' As I plug the cable into 'Kit,' the cable into the adapter, and the adapter into the EU plug, the adapter indicates that a current flows—its orange light is on. Yet the green light of 'Kit' doesn't flash, as suggested in the *User Manual*, step 72.—Why!? Is it the 'Kit,' is it the 'cable,' is the 'adapter'? Is it a question of 'time'? etc.—Before I call it a day, I check out 'help' on internet, via Google to github's 'AIY Vision Kit Setup Issue #401.'—At least one other user's 'AIY Vision Kit' has a few issues!
>
> *Vignette 3*

What does 'following the manual' mean when the manual can and does *not show how* it should be followed? The 'show-say' articulation of step 71 (see Vignette 2) offers an instructive, if simplest case. Indeed, that articulation implies the expectedly successful result of its manifest instruction ('Plug your Vision Kit into a power supply', *ibid.*), but it does not show just how that result is to be arrived at (setting aside the specification of the 'wall power supply' and 'port labelled Power' to be used, *ibid.*). In turn, Vignette 3 above describes the practical work involved in attempting to 'turn it on', and therefore to 'plug

it in'. At the same time, it documents how I try to make that practical work intelligible, in and as its consistently operable course—to begin with, to myself, as the prospective user. The described attempt was made in the vein of the kit manual, rather than in terms of the promotional video.[15]

A pragmatic orientation of this kind is made manifest in the question 'what do I *actually* need […] to have the 'Kit' up and running?' (Vignette 3; emphasis added), as well as through its practical answer *in situ*. The question marks a contrast between paraphernalia and practicality, implying a preference for the latter (as prefigured by the manual's 'minimization rule': read the 'say' part *only if* the 'show' part fails—i.e., its visually mediated following fails). The practical answer, then, displays that pragmatic orientation as it discards the scenic 'workshop' configuration (as initially staged in the promotional video) as a relevant condition for testing the 'AIY Vision Kit' *in situ* (as available to me, on my desktop, at the office, on that Wednesday afternoon, in June 2019, etc.). Instead, I assemble an alternative configuration to test the kit—a 'testable configuration' which, in addition to the assembled 'Vision Kit', includes the 'US power supply cable' and 'EU power adaptor' (as part of the minimally relevant, yet indispensably required 'experimental condition' for kit testing). Eventually, I not only attend to the intelligible configuration of the 'failing kit', but also to probing how that configuration may facilitate 'root cause' diagnosis (e.g., by singling out the assembled parts as possible items of a 'check list' laid out on my desk).[16]

CONCLUSION: 'CRITICAL INQUIRY' AND ITS REFLEXIVE GROUND(S) IN MUNDANE SPECIFICS

To my initial 'turn it on' session in early June 2019, I added four weekly sessions at the office to assemble the 'Vision Kit', and to 'make it work' in particular. 'In particular', in this case, meant not only to be able to 'turn it on' (as a first step condition), but also to 'boot it up' (as a second condition—'*kit operating condition no. 2*'). In the user manual, this second step condition (Step 72) is presented both as a visual confirmation of the preceding one, the kit's 'power connection' indicated by a 'LED light flashing', and as a software upload, whose completion would trigger a 'short tune' (*ibid.*). Yet, in practice, as no visual confirmation was given, no electric current seemed to have been flowing, rendering

[15] Hence, the 'impression of routine' may be both unavoidable and 'not groundless' (Schegloff 1986:113). Yet 'it is strategically misleading' (*ibid.*), as much for reflexive ethnography *hic et nunc* as for conversation analysis elsewhere (Schegloff's craft and concern).

[16] Noticeable inconsistency #6: 'At least one other user's 'AIY Vision Kit' has a few issues'—does it? At the time of writing (August 2020), I indeed counted up to '559 issues' (of which 124 remain 'open') regarding Google's 'AIY' kits on the software self-help site *github* (https://github.com/google/aiyprojects-raspbian/issues). In short, neither persuasive promotion nor effective instruction prevent 'shop floor problems' (e.g., Garfinkel 2002:269–270). Conversely, other users' issues might be turned into present solutions, such as github user 'Issue #176' with the 'Voice Kit'. It reads as follows: '*Hello .. I have strange issue. I bought the AIY kit from Microcenter and when I put it together using the packaged image, HAT [hardware attached on top] does not light and button does not either. I have bought three kits and its all the same. I thought it was power was the issue since I have three kits but that does not seem to be the issue. […] Its so odd .. any help I would appreciate*' (https://github.com/google/aiyprojects-raspbian/issues/176; emphasis added).

impossible the software 'boot-up' too. In the course of the four follow-up sessions, I not only figured out various 'accommodations' to deal with the pending problem (by measuring the current otherwise, using another cable, updating the software, etc.), but also discovered a series of 'resistances' induced by those very dealings (i.e., further obstacles thenceforth standing in the way of smooth kit operation). Paradoxically, every time I thought to have found an instant solution to the pending task, it turned into an additional problem to be dealt with. In short, I was confronted with a seemingly open-ended series of odd 'solution-problem pairs' (*sic*).[17]

How does 'critical inquiry' fit into the picture now? As a reflexive ethnography, this paper probed 'DIY AI' from within the technical self-instruction that it entailed—that is, in particular, through the kit assembly that it manifestly presupposed (Part I) and practically required (Part II). As an ethnomethodological topic, kit assembly was probed in its contingent circumstantiality and ordinary availability. Part I examined the 'AIY Vision Kit' video for the mundane procedures of its 'persuasive montage', didactic and conversational, thereby specifying its difference from any empirically working kit. Part II engaged us in 'DIY' kit assembly to experiment with, if not accomplish the kit's 'material production' *in situ*. We thus measured the gap between 'textual instruction' and 'practical enactment', not to mention the difference between persuasive promotion and technical work. Finally, I just noted the 'problem multiplication' that the practical enactment of textual instruction entailed, as the 'very *way* of accomplishing the task multiplied its [problematic] features' (Garfinkel 1967:26; emphasis in original). Hence, kit assembly and its 'tutorial problems' (Garfinkel 2002) manifestly do open up contrasting opportunities for critical inquiry: first, listed as 'noticeable inconsistencies' organized into as many empirically indexed footnotes (Parts I and II); now, articulated in terms of virtually indefinite 'problem multiplication' (as a current tester would not work right away, another cable would prove tricky, a software update impossible, if not unnecessary, etc.).[18]

With respect to critical inquiry into 'DIY (AI)', this emerging contrast in mundane specifics seems of particular interest, insofar as it highlights the contrasting character of its reflexive ground(s). On the one hand, the *promotional video*, insofar as it staged the 'ordinary circumstances' of the operating kit through audiovisual montage, invited us to probe inconsistencies in and through that selfsame montage. In so doing, I was led to question the 'AIY Vision Kit' as a real-worldly figure, rather than the 'DIY' ground—the workshop setting—against which it was talked and shown 'into being' (although

[17] The tentative engagement in kit assembly, as an interrelated set of technical tasks, proved similarly consequential to engaging in 'literal description': the 'assembling' itself [rather than the writing] developed the 'kit' [rather than its description] as an indefinitely 'branching texture of relevant matters' (Garfinkel 1967:26). In that sense, 'the very *way* of accomplishing the task multiplied its [problematic] features' (*ibid.*). On experiencing 'solution-problem' pairs in experimental physics, see Sormani (2014: 42–48).

[18] For the observed kind of contrast, McHoul (1994) coined a programmatic expression: 'two-way indexicality' (p. 110). As practical intelligibility may be determined *in situ*, its mundane unequivocality can also be undermined *in situ*: 'Cases in point would be some (but not all) instances involving poetry, madness, bold hypothesizing (*à la* Popper), expressionist art, […], postmodern fiction […], […] punning, irony, making *double-entendre* jokes […] and a much larger range of quotidian events' (*ibid.*, pp. 110–111).

drawing and redrawing screenshots from the promotional video invited such background inquiry). On the other hand, the longer *kit assembly* took, and the trickier it eventually got, the more it led me away from assuming the workshop setting as a reliable ground to have the kit work as a relevant figure. In other words, the advertising promise of easily getting cogent 'AI', for all its audiovisual montage, could not but turn out to be an empty one, at least in the ordinary circumstances of this 'DIY' endeavour. At this point, I cannot tell whether kit failure will turn out to be principled (as Wittgenstein quoted in the epigraph would suggest) or temporary (as the papers subsequent to this one will probe, see note 1). Hence the presently bracketed expressions ('DIY (AI)', 'reflexive ground(s)').[19]

Notwithstanding or precisely because of the open question, the paper invites practical indexation of recent reflection on critical inquiry in and across science and technology studies (STS), algorithm studies, and social and cultural studies more broadly. First, take Mirowski's (2020) discussion of the programmatic tension between casting society as a 'totality' and/or probing its 'multiplicity' in STS, and the consequences for its 'politics' of resolving that tension in one way or other (e.g., in terms of 'political economy'). How does this tension look like from within 'DIY AI' *in situ*, as questionable as the latter appears to be? Then, take Mackenzie's (2017) question—'what can critical thought […] learn from machine learning [ML]?' (p. 10)—and its ambivalent answer too: ML is deployed by 'platform-based media empires' attempting to 'capture sociality'; yet ML also produces 'knowledge of differences' (*ibid.*), a key resource for critical thinking. As reflexive 'fellow microathletes' (Sudnow 1983:93), we didn't get into the position of experiencing Mackenzie's ambivalence yet. At what level, though, is such ambivalence to be experienced? Also, would it make sense and, if so, in what sense? Finally, take Tsilipakos' (2018) argument against literary moves in moral philosophy, moves that fudge practical intelligibility and normative validity, and be it in the name of 'immanent critique' (e.g., Crary 2018). Having engaged in 'DIY (AI)' in some detail, we found its technical practice to afford us with numerous opportunities for immanent criticism—discursive and material, practical and principled. Does this experience change anything to Tsilipakos' argument? Each of the raised questions points well beyond the scope of this paper. If readers of this final paragraph have this impression too, they have come upon what I can now formulate as the paper's broader and perhaps deeper point, namely: to demonstrate how the reflexive grounding of 'critical inquiry' in mundane specifics, if paradoxically, fosters critical imagination beyond any predefined horizon (e.g., 'machine intelligence').

REFERENCES

Baccus, Melinda D. 1986. 'Multipiece Truck Wheel Accidents and Their Regulations'. In *Ethnomethodological Studies of Work*, edited by Harold Garfinkel, 20–59. London: Routledge & Kegan Paul.

Ball, Michael, and Gregory Smith. 2019. 'The Orderliness of Flow'. *Ethnographic Studies* 16: 299–314.

Bloom, Peter. 2020. *Identity, Institutions and Governance in an AI World*. Singapore: Palgrave MacMillan.

[19] In turn, 'failed repair may be politicizing, in that it leaves the problem unsolved, thus making it available to a (digital) political handling' (Strebel et al. 2019:21, note 5). Also see Stalder (2017).

Brooker, Phillip, William Dutton, and Michael Mair. 2019. 'The New Ghosts in the Machine: 'Pragmatist' AI and the Conceptual Perils of Anthropomorphic Description'. *Ethnographic Studies* 16: 272–98.

Burns, Stacy Lee. 2012. '"Lecturing's Work": A Collaborative Study with Harold Garfinkel'. *Human Studies* 35(2): 175–92.

Button, Graham, Jeff Coulter, John R. Lee, and Wes Sharrock. 1995. *Computers, Minds and Conduct*. Hoboken, NJ: Wiley.

Button, Graham, and Wes Sharrock. 1995. 'The Mundane Work of Writing and Reading Computer Programs'. In *Situated Order: Studies in the Social Organization of Talk and Embodied Activities*, edited by Paul ten Have and George Psathas, 231–58. Washington, DC: University Press of America.

Cardon, Dominique, Jean-Philippe Cointet, and Antoine Mazière. 2018. 'La revanche des neurones: L'invention des machines inductives et la controverse de l'intelligence artificielle' ['The Revenge of The Neurons']. *Réseaux* 5(211): 173–220.

Crary, Alice. 2018. 'Wittgenstein Goes to Frankfurt (and Finds Something Useful to Say)'. *Nordic Wittgenstein Review* 7(1): 7–41.

Dant, Tim. 2019. 'Inside the Bicycle: Repair Knowledge for All'. *Repair Work Ethnographies: Revisiting Breakdown. Relocating Materiality*, edited by Ignaz Strebel, Alain Bovet and Philippe Sormani, 238–312. Singapore: Springer/Palgrave Macmillan.

DiSalvo, Carl. 2012. *Adversarial Design*. Cambridge, MA: MIT Press.

Garfinkel, Harold. 1967. *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

Garfinkel, Harold. 2002. *Ethnomethodology's Program. Working Out Durkheim's Aphorism*. Lanham, MD: Rowman & Littlefield.

Garfinkel, Harold and D. Lawrence Wieder. 1992. 'Two Incommensurable, Asymmetrically Alternate Technologies of Social Analysis'. In *Text in Context: Contributions to Ethnomethodology*, edited by Graham Watson and Robert M. Seiler, 175–206. Newbury Park, CA: Sage.

Google. 2018. *Vision Kit. Do-it-yourself Intelligent Camera.* User manual, Vision Kit version 1.1. Available at: https://aiyprojects.withgoogle.com/vision (accessed 5 August 2020).

Google. 2020a. *AIY: Do-it-yourself Artificial Intelligence.* Projects website, including Promotional Videos of Vision and Voice Kit. Available at: https://aiyprojects.withgoogle.com (accessed 5 August 2020).

Google. 2020b. *Vision Kit*. Project website. Available at: https://aiyprojects.withgoogle.com/vision/ (accessed 5 August 2020).

Google. 2020c. *Voice Kit*. Project website. Available at: https://aiyprojects.withgoogle.com/voice/ (accessed 4 August 2020)

Jayyusi, Lena. 1988. 'Towards a Socio-Logic of the Film Text', *Semiotica* 68(3–4): 271–96.

Lake, Brenden M., Tomer D. Ullman, Joshua B. Tenenbaum, and Samuel J. Gershman. 2017. 'Building Machines that Learn and Think like People'. *Behaviorial and Brain Sciences* 40: 1–72.

Lynch, Michael. 1982. 'Technical Work and Critical Inquiry: Investigations in a Scientific Laboratory'. *Social Studies of Science* 12(4): 499–533.

McHoul, Alec. 1994. 'Towards a Critical Ethnomethodology'. *Theory, Culture & Society* 11: 105–26.

Mackenzie, Adrian. 2017. *Machine Learners. Archeology of a Data Practice.* Cambridge, MA: MIT Press.

Maynard, Douglas W. 1997. 'The News Delivery Sequence: Bad News and Good News in Conversational Interaction'. *Research on Language & Social Interaction* 30(2): 93–130.

Mehan, Hugh. 1979. *Learning Lessons: Social Organization in the Classroom.* Cambridge, MA: Harvard University Press.

Mirowski, Philip. 2020. 'Democracy, Expertise and the Post-Truth Era: An Inquiry into the Contemporary Politics of STS'. Draft Paper, Version 1.1. Available at: https://www.academia.edu/42682483/Democracy_Expertise_and_the_Post_Truth_Era_An_Inquiry_into_the_Contemporary_Politics_of_STS (accessed 5 August 2020).

Moore, Robert J. and Raphael Arar. 2019. *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework.* New York: ACM.

Ratto, Matt and Megan Boler. Eds. 2014. *DIY Citizenship. Critical Making and Social Media.* Cambridge, MA: MIT Press.

Rutledge, Billy. 2017. 'Introducing AIY Vision Kit: Make Devices That See'. *Google Blog 'Machine Learning'*, 30 November 2017. Available at https://www.blog.google/topics/machine-learning/introducing-aiy-vision-kit-make-devices-see/ (accessed 4 August 2020).

Schegloff, Emanuel A. 1986. 'The Routine as Achievement'. *Human Studies* 9(2): 111–51.

Schmidt, Kjeld and Joergen Bansler. 2016. 'Computational Artifacts: Interactive and Collaborative Computing as an Integral Feature of Work Practice'. Presented at *12th International Conference on the Design of Cooperative Systems*, Trento, Italy (23–27 May 2016).

Shanker, Stuart G. 1988. 'The Dawning of (Machine) Intelligence'. *Philosophica* 42(2): 93–144.

Shanker, Stuart G. 1995. *Wittgenstein's Remarks on the Foundations of AI.* London: Routledge.

Sormani, Philippe. 2014. *Respecifying Lab Ethnography. An Ethnomethodological Study of Experimental Physics.* Aldershot: Ashgate.

Sormani, Philippe. 2019. 'Ethnomethodological Analysis'. In *SAGE Research Methods Foundations: An Encyclopedia*, edited by Paul Atkinson, Sara Delamont, Alexandru Cernat, Joseph W. Sakshaug and Richard A. Williams. London: Sage.

Sormani, Philippe. 2020. 'Reenactment as Research Strategy: From Video Analysis to Performance Art and Back Again'. In *Dialogues Between Artistic Research and Science and Technology Studies*, edited by Henk Borgdorff, Peter Peters and Trevor Pinch, 184–200. London: Routledge.

Sormani, Philippe. Forthcoming a. 'Reenacting Raspberry Pi: "A Perceptron in Just a Few Lines of Python Code"'.

Sormani, Philippe. Forthcoming b. 'Bringing It All Together? Of Montage, "AI" Assemblage, and Methods—Again'.

Stalder, Felix. 2017. *The Digital Condition*. Hoboken, NJ: Wiley

Strebel, Ignaz, Alain Bovet, and Philippe Sormani, eds. 2019. *Repair Work Ethnographies: Revisiting Breakdown, Relocating Materiality*. Singapore: Springer/Palgrave MacMillan

Sudnow, David. 1983. *Pilgrim in the Microworld: Eye, Mind, and The Essence of Video Skill*. New York: Warner Books.

Techgig. 2020. '5 best AI-enabled apps that make smartphones smarter', *Techgig Coding Platform*. 20 February 2020. Available at: https://content.techgig.com/5-best-ai-enabled-apps-that-make-smartphones-smarter/articleshow/74226304.cms (accessed 28 August 2020).

Tsilipakos, Leonidas. 2018. 'Social Criticism, Moral Reasoning and the Literary Form'. *Nordic Wittgenstein Review* 7(2): 77–109.

Turner, Roy. 1970. 'Words, Utterances, Activities'. In *Understanding Everyday Life: Towards a Reconstruction of Sociological* Knowledge, edited by Jack D. Douglas, 169–87. Chicago, IL: Aldine.

Wittgenstein, Ludwig. 1960. *The Blue and Brown Books*. Oxford: Basil Blackwell.