



## Training Exercises C3 (Manipulation, Deception, and Illusion)

---

### Issue 1: *#varoufake*

Jan Böhmerrmann created a lot of confusion with Varoufake. Was he morally allowed to claim that he faked the video? Was he disallowed to do so? Give your intuition and reason to believe in your intuition.

If you want to see more check out the original clip from the show: <https://youtu.be/Vx-1LQu6mAE> (English subtitles are available)

### Issue 2: *Ought I to publish?*

Think about the following claims and their implications. Do you think that they are true?

1. Researchers are morally forbidden to publish research results that can have very drastic negative consequences.
2. That people loose trust in media is a very drastic negative consequence.
3. Open-source tools that allow for easy and convincing video manipulation can have as a consequence that people loose trust in media.
4. Deep fake should never have been published.
5. There are no countermeasures to the negative consequences of apps like Deep Fake.

### Issue 3: *Red Flag Laws*

Make a moral case for or against the following claim or a *reasonable* conditionalization thereof:

Introducing a Red Flag Law for AI is morally right.

Give a compelling argument in extended standard form that can, but does not have to, incorporate the use of moral theories. Make sure to give sufficient reason for your premises.

### Issue 4: *Beauty Filters*

Familiarize yourself with the effect that the exposure to retouched photos can have on users, especially on teenagers. You might want to take a look at what is sometimes called “Snapchat Dysmorphia”. Try to answer the following questions:

- (a) Is it morally permitted to use beauty filters for a given user?
- (b) Is it morally permitted for online services to promote the use of beauty filters?
- (c) What does this have to do with the collective action problem?

If you want to know more about the impact of social media on the wellbeing of teenagers and young adults, check out 2017's #StatusOfMind report: <https://www.rsph.org.uk/our-work/campaigns/status-of-mind.html>

### **Issue 5: *Free Choice***

- (a) If you were to explain the either-or problem of freedom to your mother, how would you do it?
- (b) Write down the argument from the either-or problem or freedom as an extended standard form.
- (c) Do you find the conclusion of the problem scary?
- (d) Do you agree with the argument or do you have an objection in mind?

### **Issue 6: *Dark Patterns***

During the video, it was said that Dark Patterns usually are bad and that employing them *ceteris paribus* is not morally permissible. Why is this true?

### **Issue 7: *Nudging***

In the lecture we very briefly discussed whether it is ok (rationally and morally) to take someone's autonomy away with nudging. Now look at the three aspects of this that probably are most relevant for your private and professional life. Argue for or against the following claims or a reasonable conditionalization thereof:

- (A) It is rational (i.e. in my best interest) to avoid being nudged by technologies.
- (B) It is morally obligatory to avoid being nudged by technologies.
- (C) It is morally permissible to implement a nudging technology.

During this process, also think about the following questions:

- (i) Where do you use nudging technologies?
- (ii) Do you try to avoid nudging and if yes, where and why?
- (iii) Do you think it makes a difference whether you purposefully decide to use a nudging technology, or whether you use it without realizing that is (or tries to be) nudging?

**Issue 8: *Electronic Whip***

Research on what is called Disneyland's "electronic whip". Where does this belong in the realm of what we discussed in the module? Does any of this sound familiar to you? In fact, the Tim-case from the Ethics exercises was inspired by this. Where are similarities and where are differences between the two cases? How do you assess the moral status of implementing either technology?

**Issue 9: *Autonomy***

Which of the definitions of autonomy do you find most compelling and why? Can you come up with a counterexample against one of them?