



# The Concept of Entrapment

Daniel J. Hill<sup>1</sup> · Stephen K. McLeod<sup>1</sup> · Attila Tanyi<sup>2</sup>

© The Author(s) 2017. This article is an open access publication

**Abstract** Our question is this: What makes an act one of entrapment? We make a standard distinction between *legal entrapment*, which is carried out by parties acting in their capacities as (or as deputies of) law-enforcement agents, and *civil entrapment*, which is not. We aim to provide a definition of entrapment that covers both and which, for reasons we explain, does not settle questions of permissibility and culpability. We explain, compare, and contrast two existing definitions of legal entrapment to commit a crime that possess this neutrality. We point out some problems with the extensional correctness of these definitions and propose a new definition that resolves these problems. We then extend our definition to provide a more general definition of entrapment, encompassing both civil and legal cases. Our definition is, we believe, closer to being extensionally correct and will, we hope, provide a clearer basis for future discussions about the ethics of entrapment than do the definitions upon which it improves.

**Keywords** Civil entrapment · Crime · Entrapment · Intentions · Legal entrapment · Proactive law enforcement

## 1 Introduction

Entrapment involves two main parties: the entrapping party and the party that the entrapping party intends to entrap. We call the first the “agent” and the second the “target”. Let these terms include groups as well as individuals. When the agent is responsible for law enforcement, or (as in the case of an informer asked by the police to

---

✉ Stephen K. McLeod  
skmcleod@liv.ac.uk;  
<http://www.liv.ac.uk/philosophy/staff/stephen-mcleod>

<sup>1</sup> Department of Philosophy, University of Liverpool, Mulberry Court, Mulberry Street, Liverpool L69 7ZY, UK

<sup>2</sup> Department of Philosophy, University of Tromsø, 9037 Tromsø, Norway

entrap) acting on behalf of someone who is, and is acting (permissibly or otherwise) in the agent's capacity as a law-enforcement agent or as the deputy of a law-enforcement agent, we are dealing with *legal entrapment*. Legal entrapment may, but need not, be *illegal* in another sense of the word "legal". When the agent is neither acting for, nor acting as, a law-enforcement officer, we are dealing with *civil entrapment*.<sup>1</sup> Civil entrapment is carried out by someone who is either not a law-enforcement officer, or the deputy of such an officer, at all, or who is but is not acting (permissibly or otherwise) in that official capacity.

An act of civil entrapment and another of legal entrapment may each involve the temptation to commit a crime of the same type. We believe that it is useful not to restrict the notion of entrapment to cases in which the agent intends that a *crime* be committed by the target.<sup>2</sup> Doing this provides theoretical utility, because like cases can be analysed in like ways (and subsequently—though we do not do so here—ethically evaluated in like ways). Accordingly, we classify acts of entrapment via the two-dimensional matrix provided in Table 1.

Type 1 = 1A + 2A = legal entrapment to commit a crime.

Type 2 = 1B + 2A = civil entrapment to commit a crime.<sup>3</sup>

Type 3 = 1B + 2B = civil moral entrapment.

Type 4 = 1A + 2B = legal moral entrapment.

We have ordered the types in this list in order of the frequency with which they appear to be discussed in the literature on entrapment, going from most to least frequent. By "moral entrapment", we mean entrapment that aims not to tempt the target to commit a crime, but rather to tempt the target to do something that is immoral, embarrassing, or socially frowned upon (measurable in part by the extent to which the target would probably not like the act to be exposed to colleagues, an employer, friends, family, or the public).<sup>4</sup>

We consider it worth including Types 3 and 4 as genuine cases of entrapment even though they do not aim at tempting the target to commit a crime. Including Type 3, for example, in the list is likely to have theoretical utility, because it will enable like analyses of like cases of legal entrapment to commit a crime by the state and entrapment, by media organizations, of targets into committing embarrassing or immoral, but non-criminal, acts that those organizations seek to expose publicly. Our approach is to recognize the diversity of—or, if you prefer, to be conceptually liberal about—the phenomenon we are studying, but to approach it, in terms of the methodology and category scheme that we employ, in a way that maximizes theoretical utility and underlying analytical unity. For the time being, however, we restrict our attention to entrapment of Type 1 so that we can make some

<sup>1</sup> The phrase "legal entrapment," as we use it, is already in currency in the literature. See, e.g., Andrew Altman and Steven Lee, "Legal Entrapment", *Philosophy & Public Affairs* 12 (1) (1983): pp. 51–69. Hock Lai Ho, "State Entrapment", *Legal Studies* 31 (1) (2011): pp. 71–95 calls legal entrapment "state entrapment" and, in common with many other authors (e.g., Kate Hofmeyr, "The Problem of Private Entrapment," *Criminal Law Review* (2006): pp. 319–336), calls civil entrapment "private entrapment". Of course, civil or private entrapment may be carried out in a highly public way (e.g., by the media). Gideon Yaffe, "'The Government Beguiled Me': the Entrapment Defense and the Problem of Private Entrapment", *Journal of Ethics & Social Philosophy* 1 (1) (2005): pp. 1–50 calls legal entrapment "government entrapment".

<sup>2</sup> In Sects. 4 and 5, we make our account of what exactly the agent intends more precise.

<sup>3</sup> To qualify footnote 1, by "private entrapment", Hofmeyr (2006) appears to mean civil entrapment to commit a crime.

<sup>4</sup> Our use of "moral" is wide. Ultimately, we think, there is an even wider notion of entrapment, corresponding to a usage of the verb "entrap" that we discuss in footnote 30. However, it is cases of the types identified in Table 1 that are of central interest in the literature and here.

**Table 1** Classification matrix for acts of entrapment

	A	B
Is the agent acting (permissibly or otherwise) in the agent's capacity as a law-enforcement agent or as the deputy of a law-enforcement agent?	Yes	No
Is the act that the agent intends the target to commit of a type that is criminal?	Yes	No

conceptual headway that we can later apply to all of the types of entrapment included in Table 1.

We recognize that the word “entrapment” may have pejorative connotations.<sup>5</sup> Nevertheless, we do not think that the proper description of an act as one of “entrapment” either does, or should, itself settle the question of the act's moral or legal permissibility.<sup>6</sup> Moreover, we are at liberty to employ a non-pejorative usage, which we shall later specify in some detail, of the term. For us, to classify a scenario as a case of entrapment is not, of itself, to say anything about its moral or legal legitimacy. (We are concerned with the denotation of the word “entrapment”, not its connotations.) On an alternative convention, which we are not following, “entrapment” is a term that is reserved for “illegitimate proactive law enforcement”.<sup>7</sup>

We distinguish three questions. First: Did entrapment *occur*? Answering this question requires an extensionally adequate *definition* of entrapment. Second: Did the *agent err* (morally or legally) in entrapping? To put it another way: When, if at all, is entrapment *permissible*? Third: Ought the *target* to be held (morally or legally) *responsible* for the target's act? That is, under what conditions, if any, is the target properly to be held *culpable* for the act that the target has been entrapped into committing?

Henceforth, we focus exclusively on the definitional question. We believe it must be answered in a rigorous and extensionally adequate manner before the other two questions can best be addressed. For reasons soon to be presented, we take it that some putative answers to this question that have featured in the literature so far have not managed to reach this goal. Our ultimate aim is to provide a definition that encompasses all the types of entrapment that feature in Table 1 and that makes some crucial improvements, in ways we explain as we go along, upon existing definitions (which cover legal entrapment only) from the literature.

<sup>5</sup> Compare Eric Colvin, “Controlled Operations, Controlled Activities and Entrapment”, *Bond Law Review* 2 (14) (2002), pp. 227–230, p. 229. Colvin holds that the term is usually pejorative.

<sup>6</sup> Compare M.L. Friedland, “Controlling Entrapment”, *University of Toronto Law Journal* 32 (1) (1982), pp. 1–30, p. 3; B. Grant Stitt and Gene G. James, “Entrapment and the Entrapment Defense: Dilemmas for a Democratic Society”, *Law and Philosophy* 3 (1) (1984): pp. 111–132, pp. 114–115. We do not intend this to prejudge the issue of whether, if a putative act is one of entrapment, the agent might thereby have a reason not to commit it. Also, we leave open the possibility that all acts of entrapment are wrong.

<sup>7</sup> This other convention is adopted by, for example, Paul M. Hughes, “What is Wrong with Entrapment?”, *Southern Journal of Philosophy* 42 (1) (2004), pp. 45–60, p. 50, by Neil Levy, “In Defence of Entrapment in Journalism (and Beyond)”, *Journal of Applied Philosophy* 19 (2) (2002), pp. 121–130, p. 122, and by Joseph A. Colquitt, “Rethinking Entrapment”, *American Criminal Law Review* 41 (4) (2004): pp. 1389–1437. It is also adopted by Andrew Ashworth, “What is Wrong with Entrapment?”, *Singapore Journal of Legal Studies* 293 (2) (1999), pp. 293–317, p. 294. Hughes leaves it open as to whether there are legitimate proactive law-enforcement scenarios, Levy argues that there are, and Ashworth works with the assumption that there are.

In Sect. 2, we explain why a *normatively thin* definition of legal entrapment (of Type 1)—that is, according to our usage, one that, of itself, does not settle the questions of permissibility or culpability—is needed.

In Sect. 3, we explain two existing normatively thin definitions of legal entrapment (of Type 1) and we point out some problems with their extensional adequacy. One of the existing definitions includes as a necessary condition of an act's being one of legal entrapment (of Type 1) that the agent arrests the target. The other includes, instead, the condition that the agent is able to observe the target committing the crime. We explain why we consider neither these conditions nor certain weakened versions of them to be necessary conditions of an act's being one of legal entrapment (of Type 1).

In Sect. 4, we provide a new definition of legal entrapment (of Type 1) that builds on the previous two but which excises what we take to be the problems with their extensional adequacy. In this definition, we replace the arrest and observability conditions with a new condition, the *traceability* condition. We explain and justify this condition. We include an additional new condition, which we also explain and justify, relating to the agent's intention to render the target vulnerable to the agent's (or a third party's) power.

In Sect. 5, we generalize our new definition of legal entrapment (of Type 1) so as to obtain a definition that encompasses all cases classified under Table 1.

In Sect. 6, we spell out some implications of our account. Section 7 includes a brief summary and some concluding remarks. In the headings of Sects. 2, 3 and 4, we use the phrase “legal entrapment”, as an abbreviation for “legal entrapment to commit a crime” (i.e., to cover cases of Type 1).

## 2 The Need for a Normatively Thin Definition of Legal Entrapment

We agree with Andrew Ashworth that the question of “what, if anything, is wrong with entrapment” requires that we should first have “a working definition of entrapment”.<sup>8</sup> We differ from Ashworth in that we believe that it is in the interests of analytical clarity, and of a prudent division of philosophical labour, that this working definition should not prejudice the answers to any normative questions about entrapment. To define entrapment, as courts in the USA tend to do, in a way that already incorporates permissibility, impermissibility, culpability, or the lack thereof is akin to defining abortion in a way that incorporates some such normative notion. It is surely one question as to what abortion *is* and another as to when, if at all, it is *permissible*. We see no good reason to treat entrapment differently. We think that to treat entrapment in the same way is likely to clarify the normative debates about entrapment and enhance our understanding of the concept.

Ashworth and others who employ a normative definition in effect treat the concept of entrapment as what we here call a “normatively thick” concept.<sup>9</sup> Let us explain this idea briefly. The concept of *killing*, for example, is normatively thin, for nothing follows merely from the description of an act as one of killing as to its legal or moral permissibility. The concept of

<sup>8</sup> Ashworth (1999, p. 295).

<sup>9</sup> A normatively thick definition of entrapment is adopted by Thomas J. Micelli, “Criminal Solicitation, Entrapment, and the Enforcement of Law”, *International Review of Law and Economics* 27 (2) (2007): pp. 258–268, p. 259: “entrapment is defined to be the unlawful arrest of a person who was not ‘predisposed’ to commit the crime in question”. For coverage of cases in the USA where the court has worked with a normatively thick definition, see Andrew Carlon, “Entrapment, Punishment, and the Sadistic State”, *Virginia Law Review* 93 (4) (2007): pp. 1081–1134, pp. 1087–1095.

*murder* is normatively thick, for to murder a person is to kill that person *unlawfully*. As noted earlier, “entrapment” might indeed carry a pejorative meaning. To describe an act as one of “entrapment” might well suggest that one takes there to be something morally wrong with it. This, however, is certainly not decisive. The question of permissibility should not be settled by simply telling us that entrapment is *by definition* impermissible. This point does not rule out the possibility that a proper definition of entrapment is non-reductive. We allow for normatively thick terms to be used in the definition—think of “crime”, which clearly has a normative element that describes an act as illegal. What we seek is a definition that does not foreclose the question of whether entrapment is permissible (whether morally or legally). Such a definition might use unreduced normative terms as well as being descriptively thick.<sup>10</sup>

While we recognize that some legal jurisdictions treat the concept of entrapment as normatively thick, we consider that it is not in the interests of philosophical inquiry into the concept and into the acceptability or otherwise of acts of entrapment to limit the concept in this way. The underlying difference between Ashworth’s approach and ours is this. His attempt to build a working definition is based on how entrapment is defined legally in jurisdictions that have an entrapment defence or in which entrapment can be grounds for stay of legal proceedings against the accused, or for the exclusion of certain evidence. We base our approach on two existing definitions from legal philosophy that have what we take to be the advantage of providing for acts to count as acts of entrapment even when those acts are not committed in jurisdictions in which the concept of entrapment is in legal currency. Their greater generality in this respect, we think, means that they have greater analytical power: that is, they will enable us to understand more than do definitions that adopt the rival, normatively thick, approach to defining entrapment.

Ashworth, after concluding that his initial search for a working definition “has not been successful”, writes that “we have identified two disparate models, one centred on the conduct of the official and the other centred on the conduct of the defendant”.<sup>11</sup> These are called the “subjective test” and the “objective test”. Andrew Carlon observes:

The two tests are phrased quite differently, but they share a structural similarity. In both, the first element is typically an essentially identical factual question of causation: Did the police induce the defendant to commit the crime? The next element in each test is a critical hypothetical: in the subjective version, the question is whether the crime would have occurred if this particular defendant had not been encouraged by the police; in the objective version, the question is whether the crime would have occurred if the particular encouragement offered by the police had been offered to a non-predisposed person.<sup>12</sup>

B. Grant Stitt and Gene G. James are right that neither test “is a test of whether entrapment *in fact* took place”; rather, they are approaches to answering the question of

<sup>10</sup> This last qualification is important because in the ethics literature, the thick/thin distinction is different. There is a well-known distinction, originating from the work of Bernard Williams, between thick and thin ethical concepts whereby an ethical concept is thick if it is *descriptively* thick, i.e., has dense and rich descriptive content. Virtue and vice concepts, such as the concepts *generous*, *cruel*, and *tactful* are considered ethically thick, whereas concepts such as *right*, *bad*, and *permissible* are taken to be ethically thin. For more on the distinction, see Pekka Väyrynen, “Thick Ethical Concepts”, in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), <http://plato.stanford.edu/archives/win2016/entries/thick-ethical-concepts/>. Our usage of the terms “thick” and “thin” is different and refers to the normative aspect, which we do not take to foreclose debate on moral/legal permissibility (which, arguably, many ethically thick as well as thin concepts do). Moreover, we remain neutral on whether the concept of entrapment is ethically thick.

<sup>11</sup> Ashworth (1999, p. 298).

<sup>12</sup> Carlon (2007, p. 1092).

culpability.<sup>13</sup> Thus, neither test should be regarded as a candidate for being a definition of entrapment. The subjective test is a test of whether an entrapment defence can be mounted, while the objective test “is a proposal for deciding whether a defendant should be tried at all” or whether, instead, the agent’s misconduct was sufficiently bad to merit dismissal of the charges.<sup>14</sup> There is an impasse between the advocates of these tests.<sup>15</sup>

Our approach, which is to define the concept of entrapment independently of these tests and independently of immediate answers to the questions of permissibility and culpability, is normatively thinner than the subjective and objective tests for it does not give such immediate answers. Even for those, unlike us, whose main concern is with the conditions under which the entrapment defence can be justified, or viably adopted, in a jurisdiction in which it is possible, analytical clarity would be enhanced, we take it, by adopting this approach. Moreover, if we are to distinguish between legitimate and illegitimate cases of legal entrapment, then we must define legal entrapment in a way that does not prejudice questions of legitimacy.<sup>16</sup>

Before we proceed, in the next section, to discuss two normatively thin definitions of legal entrapment that meet this condition, let us say a little more about what we mean when we call a definition of legal entrapment “normatively thin”. We shall do this by taking a distinction made by Hock Lai Ho and using it to illustrate our conception of normative thinness. Ho holds that, in the context of legal entrapment to commit a crime, the word “entrapment” has both a “neutral or broad sense” and a “strict or negative sense”.<sup>17</sup> On Ho’s account, an act counts as one of legal entrapment to commit a crime in the neutral sense if and only if it meets the following four conditions (which we set out using our terminology of “agent” and “target”):

1. the agent *intentionally tempts* the target;
2. the agent acts in a *covert and deceitful* manner;
3. the agent’s action is motivated, from the start, by the desire to *convict and punish* the target;
4. the agent is, or is at least acting on behalf of, a law-enforcement agent.<sup>18</sup>

Ho considers entrapment in the neutral sense to be neutral because the presence of all four of the features indicated in these conditions does not of itself imply anything about the (im)permissibility of the agent’s action. In other words, in our terms, Ho considers conditions (1)–(4) to provide a normatively thin definition of entrapment.

It might be contended, though, that the term “deceitful” is, like the term “dishonest”, normatively thick in the sense that one will describe an act as “deceitful” only if one disapproves of it morally. We have two responses to this worry. First, as we noted above, the use of a normatively unreduced (ethically thick) term in one’s definition of entrapment makes the definition normatively thick only if the use of such terms in the definition forestalls any debate on the moral/legal permissibility of entrapment itself. To see if this is

<sup>13</sup> Stitt and James (1984, p. 114). Compare Richard H. McAdams, “The Political Economy of Entrapment”, *Journal of Criminal Law and Criminology* 96 (1) (2005): pp. 107–186, p. 108; Altman and Lee (1983, p. 53).

<sup>14</sup> Stitt and James (1984, p. 114).

<sup>15</sup> For references, see Ashworth (1999), Carlon (2007), and Ronald J. Allen, “Clarifying Entrapment”, *Journal of Criminal Law and Criminology* 89 (2): pp. 407–431.

<sup>16</sup> Liat Levanon, “The Law of Police Entrapment: Critical Evaluation and Policy”, *Criminal Law Forum* 27 (2016): pp. 35–73 invokes the legitimate/illegitimate distinction.

<sup>17</sup> Ho (2011, pp. 72–73); compare Levy (2002, p. 122).

<sup>18</sup> Ho (2011, pp. 73–75).

so, we can apply an inverted version of G.E. Moore's "open question" test.<sup>19</sup> For any act,  $x$ , and any description of it as an act of  $V$ -ing that might be applied to the act, we can ask whether describing  $x$  as an act of  $V$ -ing renders the question "Is it necessarily bad/wrong to  $V$ ?" redundant. If it does not, then cases in which  $V$ -ing is not bad/wrong are not ruled out and the extent to which the description of  $x$  as an act of  $V$ -ing is normatively thick is limited. This appears to be so with Ho's definition, since it seems that we can meaningfully ask "Is this deceitful act (of entrapment) bad/wrong?". Second, even if it is difficult or impossible to define the concept of entrapment in a way that makes use of no concepts that are to some degree normatively thick, one definition can still be normatively thinner than another. Our insistence on the methodological need for a normatively thin definition of entrapment can then be interpreted as saying that we require a definition that keeps the use of concepts that are normatively thick sufficiently minimal that they do not prejudice the questions of permissibility and culpability.

Ho specifies legal entrapment to commit a crime "in the strict or negative sense" by setting out a fifth, additional, condition that an act must meet for it to count as one of legal entrapment in this sense:

5. the agent's act of intentionally tempting the target *caused* the target to commit the offence.<sup>20</sup>

Ho thinks that the causal relationship indicated in condition (5) is what we might call the "wrong-maker" for legal entrapment strictly or negatively defined. That is to say, on his account, an act that meets conditions (1)–(4) is not necessarily wrong, but one that meets conditions (1)–(5) is.

Now we reach the moral of this story. Ho's strict or negative notion of legal entrapment has in its extension, thinks Ho, only acts that are morally impermissible. This, however, is precisely a point at the level of extension, not at that of intension. It is not merely because the acts meet all five conditions that Ho considers them to be wrong acts. Rather, it is because he has independent moral reasons for considering it wrong for a law-enforcement agent, or the deputy of such an agent, to bring it about, through intentionally tempting a target, that the target commits a crime. There is certainly nothing surprising about this. We follow the standard model of ethical reasoning here, in which an ethical conclusion follows from a premise containing the relevant ethical principle (or, more broadly, the relevant ethical theory, whether principle-based or not) and an account of the ethically relevant—but otherwise ordinary, non-ethical—facts according to the principle (theory).<sup>21</sup> Our point is simply that these facts do not, in themselves, give us the ethical conclusion, but do so

<sup>19</sup> G.E. Moore, *Principia Ethica* (Cambridge: Cambridge University Press, 1903), Sect. 13.

<sup>20</sup> Ho (2011, p. 75). For now, we are not interested in the pros and cons of Ho's definition, but only in its structure. It is worth noting, nevertheless, that Ho conceives of causation in the following, counterfactual, terms: the target would not have committed a crime of the same type had the agent not put temptation in the target's way. However, there is plenty of room for doubt about whether causality can really be cashed out in terms of counterfactual conditionals. That is, it is doubtful whether the following sentences are really equivalent: "By intentionally tempting the target, the agent caused the target to commit the (type of) crime"; "If it had not been for the agent's having intentionally tempted the target, then the target would not have committed the (type of) crime". However, this doubt is irrelevant to our specific concerns in this article and our own definition of legal entrapment to commit a crime (given in Sect. 4) does not employ the notion of causation. For a survey of counterfactual accounts of causation and objections to them, see Peter Menzies, "Counterfactual Theories of Causation", in Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy* (Springer 2014 Edition), <http://plato.stanford.edu/archives/spr2014/entries/causation-counterfactual/>.

<sup>21</sup> For the standard model, see Torbjörn Tännsjö, *Understanding Ethics* (Edinburgh: Edinburgh University Press, 2nd edition, 2008), Chapter 1.

only in conjunction with the ethical principle (or theory) that designates them as ethically relevant (e.g., wrong-making).

Thus, on our account, even Ho's definition is normatively thin. This point is easily brought home. We have just set out all five of Ho's conditions: these are his candidates for ethically relevant facts.<sup>22</sup> Only one of the conditions uses a term that is arguably normatively thick ("deceitful") and, we argued, this in itself does not make the definition of entrapment normatively thick. Moreover, we have not explained, and it is not evident from condition (5), without the introduction of additional normative material (such as an ethical principle), why Ho thinks (5) describes something morally wrong.<sup>23</sup> We see, then, that a definition of legal entrapment to commit a crime that includes "wrong-making" material can still be normatively thin in the sense that it does not, without the introduction of further normative premises, prejudice questions of permissibility or culpability.

### 3 Two Definitions of Legal Entrapment

According to Stitt and James, entrapment (of Type 1) occurs (in the context of their discussion) whenever the following four conditions are all met:

1. a law-enforcement agent *plans* a particular crime;
2. the agent *induces* the target to commit it;
3. the agent *arrests* the target for having committed it;
4. *counterfactual condition*: if it were not for the agent's actions, then the token (as opposed to type of) crime would not have been committed by the target.<sup>24</sup>

On this definition, *concealment* does not, of itself, amount to entrapment. Concealment occurs when the agent has set up a means of detecting a crime and kept that means hidden from the target. For example, police officers who park their vehicles behind rows of bushes and hide behind trees with speed guns to detect speeding drivers are using concealment. *Deceptive methods* of proactive law enforcement, such as *decoy operations*, need not, of themselves, count as entrapment, for they do not necessarily meet the *induction* condition. If some police officers establish a fictional identity as an easy target for conmen, in the hope that they will catch a gang of local conmen who are preying on easy targets, but they do not do anything actively to direct the conmen to con the fictional easy target, then the officers are engaging in a decoy operation, not entrapment.

<sup>22</sup> It is an interesting further complication as to what exactly the *status* of these facts is. Ho's view that only (5) is wrong-making suggests a particularist picture, as in Jonathan Dancy, *Ethics Without Principles* (Oxford University Press, 2004), esp. Chapter 2, on which (1)–(4) serve the roles of what Dancy calls "enablers" (and perhaps also the roles of what he calls "intensifiers"). Others—who adopt the orthodox, mainstream opinion, we might say—would hold that all five facts are of the same kind and serve the same role.

<sup>23</sup> Ho (2011, pp. 77–78) provides this additional normative material. It appears ultimately to rest—although he does not say so himself—on the deontological principle that it is wrong to use a person merely as a means, and that an action that meets condition (5), in addition to the other four conditions Ho specifies, will be one in which the agent treats the target as a mere means. Ho (2011, p. 80) thinks that the intrinsic wrongness of legal entrapment "lies in the state having caused a citizen to commit a type of crime that he or she would otherwise never have done" but he has something to say about *why* this is wrong. The explanation of why it is wrong is not part of the *definition* of the concept of entrapment in Ho's narrow or strict sense.

<sup>24</sup> Stitt and James (1984, p. 114). This appeal to the token side of the type/token condition contrasts with Ho's discussion, which appeals to the type side. (See footnote 20.) With Stitt and James, we do not think that it is impossible to entrap a target into committing a token crime of a type that the target was already inclined to commit. On Ho's definition, however, this is impossible.



Regarding (2), the induction condition, we understand induction to involve the *active suggestion* to the target by the agent that the crime be committed. So, merely asking suspected drug dealers whether they have any drugs (rather than actively asking them to sell drugs) does not, on this definition, count as entrapment. There is, we take it, a distinction between presenting a target with an opportunity (e.g., by purposefully leaving a wallet on the pavement) and encouraging the target (whether verbally or otherwise) to use that opportunity (e.g., by saying, with the intention to encourage, “You could steal that!”). Decoy operations and other deceptive methods of proactive law enforcement fall in the former category, while legal entrapment belongs to the latter.<sup>25</sup> Definitions of entrapment can be unclear about the distinction between presenting someone with an opportunity and encouraging that person to use it. As in Ho’s definition earlier or Gerald Dworkin’s upcoming definition, they often use ambiguous terms such as “temptation” or “enticement”.

Let us consider Dworkin’s definition of legal entrapment (of Type 1).<sup>26</sup> Dworkin holds that *proactive law enforcement* occurs whenever a law-enforcement agent (presumably, acting in that agent’s official capacity):

1. uses *deception*;
2. to *produce* the performance of a crime;
3. in circumstances in which the agent *can observe* the performance of the crime.

Dworkin regards all cases of legal entrapment (of Type 1) as cases of proactive law enforcement that meet the following additional conditions:

4. the agent *procures* the crime (by solicitation, persuasion, or enticement);
5. *counterfactual condition*: the target would not have committed the particular crime but for the agent’s having procured it.<sup>27</sup>

As in Ho’s definition, Dworkin’s definition includes a deception condition. Stitt and James do not include a deception condition. Our own view is that deception is a widespread, but probably contingent, feature of legal entrapment to commit a crime. For example, it is arguable that a plain-clothes detective who approaches a suspect and asks to buy drugs from the suspect is not necessarily *deceiving* the suspect. A detective may perfectly well solicit a crime in a manner that involves only withholding true information (e.g., that the detective is a detective, rather than a genuine customer), rather than giving false information. If deception involves giving false information, rather than merely withholding relevant true information, then it is arguably inessential to legal entrapment to commit a crime. Even if deception can involve an agent’s withholding relevant true information in a context in which the hearer would reasonably expect (as a matter of linguistic pragmatics) to be provided with it, and when the agent’s intention is that the hearer will be misled, it would still seem to be inessential to entrapment. We prefer not to dwell here on the nature of deception, which raises philosophical questions that could occupy us for some considerable time and, in Sect. 4, we follow Stitt and James in not including a deception condition in our definition.

<sup>25</sup> For more on the distinction between legal entrapment and proactive law enforcement, see Stitt and James (1984, p. 126) and Gerald Dworkin, “The Serpent Beguiled me and I Did Eat: Entrapment and the Creation of Crime”, *Law and Philosophy* 4 (1) (1985): pp. 17–39.

<sup>26</sup> Dworkin (1985).

<sup>27</sup> Dworkin (1985, pp. 17, 21). Dworkin, like Stitt and James and unlike Ho, appeals to the token side of the type/token distinction. (See footnotes 20 and 24.) On the distinction between proactive law enforcement and legal entrapment, and on their relationship, compare Ashworth (1999).

While Stitt and James work with an *arrest* condition, Dworkin works with an *observability* condition. We consider both conditions to be too strong.

The arrest condition is too strong because an agent intent on entrapment could succeed in inducing the target to commit the crime and the target could die immediately after, or while, committing it. The failure to arrest would not seem to amount to its not having been the case that entrapment occurred. One might weaken the arrest condition so that it is the agent's having an *intention to arrest*, rather than actually making an arrest, that counts. Against this, consider a third party who is a law-enforcement agent but who is not party to the satisfaction of conditions (1) and (2) of the definition by Stitt and James. Assume that this third party has no prior knowledge of the situation, including the entrapping agent's role in it and the fact that the target is a target, prior to the crime being committed. Assume further that condition (4) of the definition by Stitt and James is met. If a law-enforcement agent, A, who has a target in mind, merely intends that a third party, B, like the one just described, who is not working with A as an entrapping party, should make an arrest, then this does not preclude A's act from counting as one of entrapment. In response, a further weakening would say that it is not the intention to arrest that matters, but the intention that the target be arrested (by someone or other). If, however, a corrupt law-enforcement agent threatens to arrest in order to blackmail, with no intention that any arrest be made, then, we take it, this does not preclude the act on the part of the corrupt agent that enabled the blackmail from counting as an act of entrapment. (As we noted in Sect. 1, even an illegal act can count as an act of Type 1).

Here is a story to show that Dworkin's observability condition is too strong. Suppose that an agent has already deceived the target, planned for the target to commit the crime, induced the target to commit it, and that the agent intends to observe it. While travelling to the intended crime scene, the agent is involved in a car accident. As a result, the agent arrives an hour after the crime happened, and there were no witnesses. Nevertheless, there is evidence that is sufficient to secure the arrest of the target because, while he was running away, fibres from the target's clothing, containing his DNA, were left on a screw that was jutting out of a doorway. The agent had planned that the crime should be *detectable*, because the agent intended to witness it. While the agent did not intend the manner in which the actual crime was detectable—the snagging of the fibres on the screw was purely an accident—the agent did intend *that* the crime be detectable. This is clearly a successful case of entrapment, even though no one actually witnessed the crime and, in jurisdictions such as the USA and its constituent states, in which an entrapment defence is in principle possible, an entrapment defence could certainly be mounted by the target's counsel. In a jurisdiction in which there is no entrapment defence but in which entrapment can serve as a reason for stay of legal proceedings, or as a reason to rule inadmissible certain evidence, against a defendant,<sup>28</sup> the case could likewise be treated as offering such grounds.<sup>29</sup>

<sup>28</sup> For details and reference to cases, see Mike Redmayne, "Exploring Entrapment" in Lucia Zedner and Julian V. Roberts (eds.), *Principles and Values in Criminal Law and Criminal Justice: Essays in Honour of Andrew Ashworth* (Oxford: Oxford University Press, 2012), pp. 157–170, pp. 158–159. Picture an identical story to the one just told but in which there is no incriminating evidence. On our account, this is not thereby precluded from being a case of entrapment. Of course, the scenario is not a success from the agent's point of view: the agent, through lack of incriminating evidence, does not obtain the power (either for the agent or for a third party) to press charges against the target. We take it that whether entrapment has occurred is one thing; whether the agent (or anyone) *knows* that it has is another.

<sup>29</sup> Fiona Leverick and Findlay Stark, "How do You Solve a Problem Like Entrapment? *Jones and Doyle v HM Advocate*", *Edinburgh Law Review* 14 (3) (2010): pp. 467–472, note (p. 467) that: "The issue of how a claim of entrapment should be dealt with by the criminal courts has divided the international legal community, with different jurisdictions regarding it as a substantive defence, a matter that should lead to a stay in proceedings or a matter that should lead to the exclusion of the evidence obtained". For further details of

Moreover, a law-enforcement agent could entrap a target into committing a crime that the agent does not even intend anyone but the target to observe. Suppose, for example, that the agent has good reason to believe that the target is susceptible to suggestion. The agent encourages the target to take up a criminal opportunity that would not otherwise have been available to the target. The agent does not observe, does not set out to observe, and does not intend that a third party should observe, the crime being committed. What the agent cares about is securing a prosecution and the agent is willing to bet that, when accused of having committed the crime, the target will confess even though there were no witnesses. The upshot of this example is that, in order to entrap a target, a law-enforcement agent must intend that there is either sufficient non-verbal evidence, or sufficient testimony (including, where admissible, the target's own confession), to justify bringing charges against the target. Observation can be a means to this, but neither observation nor intended observation is required.

These two cases show that neither a version of the arrest condition nor a version of the observability condition is necessary to entrapment. Instead, they suggest something weaker: that the crime be *traceable* to the target either, as in the first case, by being detectable (by a party other than the target) or, as in the second case, via testimony (including the target's confession). In short, entrapment requires that the crime be *traceable to the target by evidence that would link the target to the act*.

Certain cases of collusion that have no element of entrapment illustrate a drawback of Dworkin's definition and bring out another crucial feature of legal entrapment (of Type 1). Consider the following story. Officer A is a corrupt police officer who is involved in extorting protection money from local entrepreneurs. He deceives his colleague, Officer B, into believing that the risks of being caught for doing this are far lower than is actually the case. Through this and other procurement techniques, A persuades B to accompany A in approaching the owner of a local business, not previously known to B, to demand protection money. A happens to have no intention to have B exposed, prosecuted, or blackmailed. Here, all five of Dworkin's conditions are met, and A has duped B into committing a crime. The officers have colluded in committing the crime. It is the agent's reasons for acting that mark cases of entrapment out from cases like the action of Officer A. When an agent entraps a target, we suggest, the criminal act is not the crucial trap into which the agent is trying to lure the target. Rather, the crucial trap is for the target to fall into being vulnerable, owing to the target's involvement in the criminal act, to the agent's (or a third party's) power to have the target prosecuted, exposed, or threatened with prosecution or exposure.<sup>30</sup> Therefore, when we define legal entrapment in the next section, we include a clause to cover this aspect of entrapment scenarios.

---

Footnote 29 continued

this diversity, with references to cases, see Ho (2011, p. 71). We think this diversity provides an additional reason to seek an underlying definition of entrapment that is so normatively thin as to be able to feature neutrally in this debate.

<sup>30</sup> The *Oxford English Dictionary* entry for "entrap" includes: "To catch in or as in a trap. Also *transf.* and *fig.* to bring unawares into a position of difficulty or danger; to bring (a person) into one's power by artifice"; "To beguile, bring by artifice *to or into*". ("entrap, v.1". *OED Online*. December 2016. Oxford University Press. Accessed February 24, 2017). In this article, it is the agent's bringing the target into the agent's power, by artifice (specifically, by procurement through incitement, solicitation, or persuasion), that is relevant. In a legal context, the point of procuring the crime is to secure the power to press charges against the person committing it. Entrapment merely in the sense of beguiling is what Officer A does to Officer B. While beguiling is not what we are primarily interested in, we say a little more about the distinction between mere beguiling and entrapment (as here understood) in Sect. 5. The entry for "entrapment" includes: "*Law*. A method of criminal investigation in which the police instigate, initiate, or encourage the commission of a

## 4 Legal Entrapment: A New Definition

Legal entrapment to commit a crime occurs whenever:

1. a law-enforcement agent (or the agent's deputy), acting in an official capacity as (or as a deputy of) a law-enforcement agent, *plans* that the target commit an act;
2. the planned act is of a type that is criminal;
3. the agent *procures* the act (by solicitation, persuasion, or incitement);
4. the agent intends that the target's act should, in principle, be *traceable* to the target either by being *detectable* (by a party other than the target) or via *testimony* (including the target's confession), that is, by *evidence that would link the target to the act*;
5. in procuring the act, the agent intends to be enabled, or intends that a third party should be enabled, *to prosecute or to expose the target* for having committed the act.<sup>31</sup>

Conditions (1) and (2) adapt the first condition given by Stitt and James and, for our analytical convenience, break it into two conditions. Condition (3) is Dworkin's procurement condition, but with "enticement" replaced (due to its vagueness in respect of whether the agent actively suggests to the target that the act be committed) by "incitement". While we could say a lot about what solicitation, persuasion, and incitement are, we settle for a few brief comments. We inherit from Dworkin, make explicit, and explain, a restriction on what counts as procurement. Certain types of compulsion, such as inducing a person to commit an act by chemical means, do not fall under condition (3). When an agent procures a criminal act, the agent does not merely influence the target's will: rather, the agent influences the target's will in a certain way. The manner in which the target's will is influenced involves responsiveness to the *content* of the agent's speech act(s) or other communicative act(s). Unlike chemical induction, solicitation, persuasion, and incitement involve tempting the target, through communicative acts, to commit the criminal act (rather than simply bringing about the criminal act). The first clause of Ho's definition (set out in Sect. 2 above) correctly identifies the agent's intention to tempt the target as an element of acts of entrapment. In order to maintain a distinction between decoy operations and cases of entrapment, we specifically limit the type of temptation involved to active temptation via a communicative act or acts in which the content of the communicative act(s) inclines the target, by its influence upon the target's will, to commit an act of a criminal type. Procurement is distinct from causation. When the agent procures the target's act by persuasion, the agent influences the target's will through reason, rhetoric, or both. Solicitation and incitement leave room for the agent to have exercised less influence on the target's will, because, unlike persuasion, they do not necessarily suggest an initial reluctance or resistance on the part of the target towards committing the criminal act. A separate notable point is that there is a distinction between *enabling* an act and the stronger notion of the *procurement* of the act as specified by condition (3). For example, if an agent provides a target with the means to commit an act of a criminal type, such as a weapon, a stash of drugs, or some counterfeiting equipment, but the agent does not solicit, persuade,

---

Footnote 30 continued

crime by a suspected offender in order to secure his or her arrest; the result of such action" ("entrapment, n". *OED Online*. December 2016. Oxford University Press. Accessed February 24, 2017). Our main concern is with legal entrapment (of Type 1) and other cases of entrapment that are structurally similar, in ways Sect. 6 makes evident, to it.

<sup>31</sup> We intend condition (5) to include blackmail cases in which the agent intends not that the target will be prosecuted or exposed but that the target will be placed under threat of prosecution or exposure.

or incite the target to commit the act, then condition (3) has not been met. Conditions (4) and (5) we take to improve upon the arrest condition of Stitt and James, Dworkin's observability condition, and Ho's condition about the agent's desire to convict and punish the target.

Condition (4), the intended-traceability condition, is weaker than the arrest condition and the observability condition but—on the assumption that to intend to arrest or observe is, thereby, to intend to have evidence that will link the target to the crime—encompasses both the intention to arrest and the intention to observe. Condition (5) specifies the proper sense of what it is to be trapped as a target in entrapment cases.

One might ask whether we indeed need these two conditions to account for cases of entrapment. We think we do. When these two conditions are not fulfilled, what we have is a case of beguiling, or, if you prefer, of entrapment in a wider, colloquial, sense, rather than of the kind of legal entrapment that we are here attempting to define and which is of principal interest to the law (i.e., entrapment of Type 1).<sup>32</sup> For example, if an agent merely wants to make the target feel guilty, or to put the target in a position in which the agent has some power over the target other than the power to prosecute, expose, or to threaten to do so, then the act is one of mere beguilement rather than one of the kind of entrapment that mainly interests us. We take it that cases of mere beguilement are philosophically interesting and that they raise similar issues, concerning, for example, the ethics of subjecting others to temptation, as do the cases of entrapment that mainly interest us.

We think that our procurement condition renders the inclusion of a counterfactual condition unnecessary. Recall that we construe procurement as the actual influencing, via the content of a communicative act (or series of such acts), of the target's will. On our account, if the target's will was not so influenced, then entrapment has not occurred (whether or not the target actually commits the crime). What matters is not whether the target would have committed the crime had the target not been so influenced, but whether the target actually was so influenced. If the target would have committed the act anyway, e.g., because a seasoned criminal was waiting to procure it and to collude with the target to commit it, but was scared off when seeing the agent approach the target, then, we take it, this does not preclude what actually happened from being a case of entrapment. In contrast, on the counterfactual conditions discussed above, the scenario just described does not count as a case of entrapment because the target would have committed the act even if the agent had not procured it.

## 5 Entrapment: A New Definition

The definitions discussed so far attempt to define only legal entrapment to commit a crime. Entrapment may be defined more generally, to encompass all four types identified in Table 1, by dropping the reference to law-enforcement agents and widening the class of induced acts. Thus, with those amendments to conditions (1) and (2), entrapment occurs whenever:

1. an agent *plans* that the target commit an act;
2. the planned act is of a type that is criminal, immoral, embarrassing, or socially frowned upon (measurable in part by the extent to which the target would probably not like the act to be exposed to colleagues, an employer, friends, family, or the public);

<sup>32</sup> See footnote 30.

3. the agent *procures* the act (by solicitation, persuasion, or incitement);
4. the agent intends that the target's act should, in principle, be *traceable* to the target either by being *detectable* (by a party other than the target) or via *testimony* (including the target's confession), that is, by *evidence that would link the target to the act*;
5. in procuring the act, the agent intends to be enabled, or intends that a third party should be enabled, *to prosecute or to expose the target* for having committed the act.

As in the case of legal entrapment (of Type 1), it has sometimes been put to us by our discussants that conditions (4) and (5) are unnecessary. Here is a putative counter-example to these conditions. Consider a mafioso who takes a hanger-on out on a kneecapping job, deceiving the hanger-on into thinking that he will just be the lookout, but then manipulating the hanger-on into committing the crime himself. The mafioso may do this with no intention of traceability, but instead intending merely to establish a psychological tie that binds the hanger-on more firmly to the gang. Some of our discussants have regarded this as a case of entrapment. Our response is the same as before. Given that condition (4) is not met (also, in relation to condition (5), that the kind of power the mafioso seeks over the hanger-on is a psychological hold, rather than a power to expose), we have a case that counts as beguilement (or, if you prefer, entrapment as more liberally understood) but not one of entrapment in the sense that is mainly of interest to us and that is most relevant to legal and civil sting operations.<sup>33</sup> Here, the aim is not merely to beguile, but to subject the target to prosecution, exposure, or the threat of such. It is the vulnerable position into which the target is put, in respect of the agent's (or a third party's) power to prosecute or expose the target, not the target's act itself, that is the trap.

## 6 Consequences of Our Account

We should like to draw attention to some important consequences, ethical as well as legal, of our new definitions. First, note that on our definition, civil entrapment to commit a crime is a form of, but at the same time distinct from, criminal solicitation: every act of civil entrapment to commit a crime is one of criminal solicitation, but not vice versa.<sup>34</sup>

Secondly, on our account, whilst legal entrapment to commit a crime can be a means of proactive law enforcement, it need not be. Dworkin considers all cases of legal entrapment to commit a crime to be cases of proactive law enforcement that meet his procurement condition and his counterfactual condition. We speculate that Dworkin does this because he has in his mind scenarios in which an entrapment defence might be mounted in the USA. In such scenarios, the target will indeed have ended up in court as a result of proactive law enforcement. We think that Dworkin has confused a *functional role* that legal entrapment to commit a crime can perform (namely, as a method of proactive law enforcement) with what legal entrapment to commit a crime *is*. On our definition, a scenario in which law-enforcement agents intend to blackmail, rather than to enforce the law, can still count as a case of entrapment (including legal entrapment to commit a crime). Accordingly, while all cases of entrapment that are intended to result in the target's being charged with a crime are attempts at proactive law enforcement, not all cases of legal entrapment to commit a crime are cases of proactive law enforcement.

<sup>33</sup> If necessary, see again footnote 30.

<sup>34</sup> Contrast *United States v. Manzella*, 791 F.2d 1263, 1269 (7th Cir. 1986), which takes "private entrapment" and "criminal solicitation" to name the same thing.

Thirdly, as we noted right at the start, for us, as for Stitt and James and for Dworkin, to classify a scenario as a case of entrapment is not, of itself, to say anything about its moral or legal permissibility. The subjective and objective tests for entrapment are, as Stitt and James note, different accounts not of when entrapment *occurs*, but of when entrapped persons should be held (criminally) *responsible* for their offences.<sup>35</sup> By the lights of the subjective test, the entrapped person is to be held responsible for the offence only if the offence exhibits the *predisposition* to commit it. Whether an entrapped person can properly be judged so predisposed will partly be a result of the degree of inducement or enticement that was offered by the law-enforcement agent. Inducement and predisposition are correlated, both in fact and in law. Other things being equal, the greater the degree of inducement or enticement has to be before the person succumbs to temptation and commits the offence, the less the person can properly be held to have been predisposed to commit the offence. According to the objective test, the entrapped person is to be held responsible only if the probability that “a hypothetical law-abiding citizen” would, in the circumstances, have committed the crime falls below a certain threshold.<sup>36</sup> By the lights of the objective test, the predisposition or otherwise of the target is irrelevant. While the target’s observed behaviour is plausibly an index of predisposition, what matters legally is the extent to which the target’s observed behaviour diverges from that of the hypothetical law-abiding citizen. We now formulate the two tests in a way that makes them applicable to all of the types of entrapment identified in Table 1. In the case of the objective test, our formulation alludes to the planning and procurement conditions set out above.

*Subjective test.* An entrapped person is (morally or criminally, depending on the case) *culpable* only if the act that the person has been entrapped into committing exhibits the person’s predisposition to perform such acts.

*Objective test.* An entrapped person is (morally or criminally, depending on the case) *culpable* only if the act that the person has been entrapped into committing is one that a hypothetical ethically upstanding or, in the legal case, law-abiding, person would be unlikely to have been led, in the circumstances of the agent’s having planned and attempted to procure the act, to have attempted.

The hypothetical ethically upstanding person (and likewise, the hypothetically law-abiding person) cannot be conceived of as a paragon of virtue, as someone with unusually strong powers of resistance to temptation, or as someone who tends towards actions and omissions that are supererogatory (i.e., admirable, but beyond the call of duty). Rather, this person has to be thought of as having (on some measure) about the average degree of virtue either typical of, or expected in, the population.

## 7 Conclusion

We have explained why, in our view, a normatively thin definition of the concept of entrapment is desirable. We have examined two normatively thin definitions of legal entrapment (of Type 1), those of Dworkin and of Stitt and James, and found them wanting. We have also proposed what we take to be a new and improved definition, though we do not say that it is entirely adequate. We have extended this definition to provide a new

<sup>35</sup> Stitt and James (1984, p. 114).

<sup>36</sup> Dworkin (1985, p. 22).

general definition that we intend to cover all the sorts of entrapment that feature in Table 1. Of course, our new general definition does not prejudge the moral issues. We intend to write separate work, using the analytical work we have done here as a foundation for the discussion, to address the question of whether entrapment is ever ethically permissible.

**Acknowledgements** We thank the participants in the conference “Public Standards, Ethics and Entrapment”. The conference took place at the University of Liverpool on 19th May 2016 with support from the University of Liverpool’s Interdisciplinary Networking Fund. We wrote an early version of some of this material for it. We are grateful to members of audiences at the following conferences, at which we have presented some of this material: VIII Braga Meeting on Ethics and Political Philosophy (University of Minho), the Society for Applied Philosophy Annual Conference 2017 (University of Copenhagen) and the Mind Association/Aristotelian Society Joint Session 2017 (University of Edinburgh). For comments on earlier drafts, we thank Thomas Atkinson, Ian Dunbar, Matthew Hart, Douglas Husak, Gregory Miller, Thomas Schramme, and two anonymous referees.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.