RULE-FOLLOWING, RULE-BREAKING AND CONDITIONS FOR MORAL RESPONSIBILITY IN THE CASE OF CHERNOBYL NUCLEAR DISASTER

ISTVÁN DANKA[1]

Department of Philosophy and History of Science,
Faculty of Economic and Social Sciences,
Budapest University of Technology and Economics,
Műegyetem rkp. 3., H-1111 Budapest, Hungary
e-mail: danka.istvan@gtk.bme.hu

JÁNOS TANÁCS[2]

Department of Argumentation and Marketing Research,
Faculty of Economics,
Eötvös Loránd University,
Egyetem tér 1-3., H-1053 Budapest, Hungary
e-mail: tanacs@gtk.elte.hu

ABSTRACT

In this paper, we analyse conditions of responsibility in a rule-following framework as applied to the 1986 Chernobyl nuclear disaster. First, based on historical documents and reports, we

[1] Márta Fehér is a role model for all of us who knew her; and perhaps the only person I have ever known in the profession who has been unanimously admired by generations of her students and the students of her students. I wish I could call her a master of mine.

[2] I am grateful to Márta Fehér, who helped me start my career as a researcher in the history of science: Márta helped me realise that I have a sense of appreciation for historical details. I pay tribute to her with this historical-fact based paper on the philosophy of technology.

describe the events leading to the accident with an emphasis on rule violations done by the operators of the reactor. Second, we investigate what INSAG-7, the most accurate descriptive report on the accident, implies regarding the responsibility of operators for the disaster. Third, we introduce a distinction between formal (or legal) and moral responsibility, claiming that while the first applies to blind, command-based rule-following situations, moral responsibility can occur only when blind rule-following is given up or, more precisely, broken.

Then we shall investigate two conditions of moral responsibility: the control (or freedom) condition and the epistemic condition. On the one hand, in a rule-following framework, the control condition can be satisfied only if agents act against orders. Following the regulations, (almost) whatever operators do, doing so prevents them from being responsible for their acts. On the other hand, the epistemic condition requires agents to be in an epistemically favourable position that includes knowledge of the possible consequences of their acts. On these grounds, operators cannot be taken as responsible for the Chernobyl disaster because to their case, a disjunction applies: either the control condition or the epistemic condition cannot be met, and neither can be satisfied in a blind rule-following framework which was applied to their case.

1. Rule-violations in the Case of Chernobyl Disaster

On the 26th of April 1986, two consecutive explosions happened at Unit 4 of the Chernobyl Nuclear Power Plant (NPP) in Ukraine, near Kyiv. Due to these explosions, the reactor of Unit 4 was essentially destroyed, and a very large quantity of radioactive substances was released into the environment. Up till now, it is considered to be one of the two most serious nuclear accidents in history (the other was the Fukushima accident in 2011).

The International Nuclear Safety Advisory Group (INSAG) of the International Atomic Energy Agency (IAEA) investigated the circumstances of the disaster and published a report about the reconstructed causes of the accident in 1986. The report, INSAG-1, suggested that "the primary cause of the accident was an extremely improbable combination of violations by the unit personnel of operating instructions and procedures" (INSAG-7 1992: 30). In 1992, the IAEA published a second report about the Chernobyl Nuclear Power Plant accident and re-evaluated the statements of the first version in many ways. Currently, this second report, widely referred to as INSAG-7, is the officially accepted historical reconstruction of the events leading to the accident. This second report re-identified the causes of the accident as well as redistributed the personal and

organisational responsibilities. By the time of the second report, some very serious design flaws came to light. The second report identified these defects as the root causes of the accident.[3] In parallel, there was also a change in the evaluation of the anomalies committed by the operators and the actions carried out by them. So, the INSAG-7 report essentially took the burden of personal responsibility off the shoulders of operators (INSAG-7 1992: 24, Danka and Tanács 2021). Since there was a serious change of position concerning the responsibility of operators by the time of the second report, this shift of standpoint had to be understood and articulated well, so the anomalies committed by the operators, the actions carried out by them, and the interpretation of the compliancy of their behaviour from the viewpoint of violations remained in the focus of the second report.

Originally, the IAEA in the first report identified six violations of the instructions, rules and procedures committed by the operating personnel as major causes of the accident (INSAG-7 1992: 17). Later, in 1992, in the second report, the IAEA accepted only one violation which played a direct role in the accident out of the formerly identified six. This infringement was connected to the "operation with too low an ORM" (INSAG-7 1992: 18). ORM meant "Operating Reactivity Margin" and was defined as the extra reactivity that would arise if all control and safety rods were withdrawn from the reactor core and was expressed in terms of the number of "equivalent control rods of nominal worth remaining within the core" (INSAG-7 1992: 6). So ORM is a multiple of the total reactivity controlled by a standard rod. Section 9 of the Operating Procedures of Unit 4 nailed down that "when the ORM falls to 15 rods, the reactor should be shut down immediately" (INSAG-7 1992: 79). This value of 15 was the permissible minimum of ORM. During the 24-hour preparation period for a turbogenerator test which finally led to the accident, the operators violated the regulation of the minimum required operating reactivity margin several times. From the four shifts which worked during the 24-hour period before the accident, at least two committed the

[3] The notion of root causes is sometimes understood as "causes that, in an important sense, do not have further causes but, rather, conclude the explanation" (Coren 2019: 316). We use the term differently, putting it in contrast with contributory causes. A root cause is the primary cause, which directly causes an event on the condition that secondary, contributory causes also interplay. Contributory causes can be characterised as a set of unnecessary and insufficient causes of an event, so that their set, along with the root cause, is a sufficient (but unnecessary) cause of the event. The root cause, in contrast, is a necessary but insufficient cause, which gets to be sufficient if it co-occurs with any alternative set of contributory causes.

(f)actual violations of minimum permissible ORM, and the two other shifts committed either (f)actual or negligent infringements, depending on different historical interpretations of some uncertain facts discussed below (Danka and Tanács 2021).[4] Operating below the allowed absolute minimum of 15 rods meant that 196 rods or more of the total 211 were completely withdrawn and that many fully withdrawn control and safety rods could introduce a serious amount of positive reactivity by inserting of them at the time of the accident at 01:23 on the 26th of April 1986. The great number of fully withdrawn control and safety rods was probably a decisive contributory factor to the accident (INSAG-7 1992: 23).

According to INSAG-7, with one exception of this infringement of ORM, the violations identified in the first report were either not actually infractions in the absence of the relevant provisions (INSAG-7 1992: 24; Szathmáry and Aszódi 2005: 94), or had no role and significance in causing the accident because their observance would not have prevented the accident, nor would have reduced its severity. However, INSAG-7 claimed no longer that ORM had played the role of a root cause: violation concerning ORM became a decisive, but merely contributory factor in the accident.

Still, INSAG-7 suggested that there were two further factual or potential violations, which were neither root causes nor contributory causes of the accident, but which are nonetheless important to identify, investigate and analyse since either of them could just as well have caused the initiating event or could have been a contributory factor to an almost identical accident (INSAG-7 1992: 13-14). First, there was a violation concerning the flow rate of the main coolant pumps (INSAG-7 1992: 7-8, 18), and second, another questionable infringement case. In this latter case, the fact of the violation depends on what kind of power was meant in the Operating Procedures. Let's see these infringements in detail.

The RBMK-1000 type reactor of Chernobyl included two independent primary coolant loops, and

[4] In contrast with actual or factual rule-breaking (which are two names for the same phenomenon), when an agent directly and explicitly violates a rule, negligent rule-breaking happens if the agent does not obtain all information relevant for following the rule, and hence she, according to her best knowledge, does not do anything wrong directly, but she misses to do something that is to be done. In other words, she does not know some of the conditions of her acts or behaviour that must be satisfied and hence she does not know that she is expected to follow a rule which she misses to follow. In some interpretations, operators in Chernobyl were negligent about ORM values repeatedly, and it resulted in their negligent violation of those rules when the ORM value went below 15 without their knowing that.

each loop had four primary coolant pumps. During normal operation, six of these eight main coolant pumps were used, and two of them were on standby mode. These two standby coolant pumps were backups in case if one or two of the six would be out of order. Each pump has a capacity ranging from the flow rate 5500 to 12000 m^3/h. During normal startup or shutdown operations, as was the case when the preparation happened before the accident, the flow rate was reduced from the normal rate of 8000 m^3/h per pump to the range of 6000-7000 m^3/h. In parallel, fewer pumps than the normal six were used during the low power phases of startup and shutdown. These restrictions were to prevent cavitation in the pumps. The phenomenon of cavitation is dangerous since it could cease to circulate the coolant. The improper cooling of the reactor in the case of the RBMK-1000 type of Chernobyl could lead to a power excursion and at last to the damage of the core.

Although no provision prohibited the connection of all eight main coolant pumps to the reactor at any power level, the Operating Procedures limited the capacity of each main coolant pump to 6500-7000 m^3/h in order to prevent cavitation at low power levels. When the Kyiv power grid controller gave permission to disconnect Unit 4 from the power grid at 23:10 on the 25th of April, a rapid power reduction started from the power level of 1500-1600 MW. At 00:05 on the 26th of April 1986, the reactor thermal power was as low as 720 MW. At 01:03 and 01:07, twenty minutes before the accident, in accordance with the testing programme, the two backup main coolant pumps were switched on to cool down the reactor. So, all eight main coolant pumps were in operation at full flow not much before the accident, and "it appears that flow rates in several exceeded prescribed values" at such a low power level (INSAG-7 1992: 18). The flow rates of some main coolant pumps reached 7500 m^3/h, exceeding the prescribed value of 7000 m^3/h at this time. Therefore, these excessive flow rates violated procedures.

The other potential violation of procedures concerns the interpretation of the power type stipulated in the regulations. The problem is that the regulation had been meant to be clear and precise, but turned out to be ambiguous. Let's see the question in detail.

As we have seen above, the operators certainly violated the provision concerning the absolute minimum ORM several times when the ORM value was definitely less than 15 rods. But a more specific provision may also have been violated. This provision concerned the so-called short-term shutdown of the reactor and was related to the event of an unplanned drop in power at 00:28 on the 26th of April. At that time, an hour before the accident, the staff lost control of the reactor and an

unplanned power drop occurred. Due to this unintended power reduction, the thermal power of the reactor fell to 30 MW, while the neutron power dropped to zero. After a 4–5-minute pause, a power raise was initiated by the operator personnel. A question arises whether this event should be considered as an (unintentional) short term shutdown.

According to the Operating Procedures, reactor power rise "after a short term shutdown when there has been no iodine poisoning of the reactor is permitted only if there is the required operational reactivity margin, which is determined from the margin prior to the shutdown of the reactor" (INSAG-7 1992: 80). A table was attached to his requirement in order to determine the required operational reactivity margin as a function of the power level at which the reactor had been operating prior to shutdown (INSAG-7 1992: 80). According to the table, when the reactor, prior to a shutdown, was operated below 50% of the nominal power level, at least 30 rods were required during a power rise. The nominal power of the reactor of Unit 4 of Chernobyl was 3200 MW, and before the unplanned power reduction at 00:28, the thermal power of the reactor was equal to 720 MW at 00:05 on the 26th of April. Considering that prior to the unplanned power drop the reactor was operated at 22.5% of the nominal power level, it was true that a minimum of 30 control rods was required during a power rise initiated after a short-term shutdown.

So the question is whether the unintended power reduction at 00:28 followed by a pause lasting 4-5 minutes after the drop is a short-term shutdown or not. It depends on the definition of the "short-term shutdown". The Operating Procedures defined the short-term shutdown of the reactor as "a reactor power reduction to zero without cooling of the MFCC" (INSAG-7 1992: 74), where MFCC means the "multipass forced circulation circuit". The problem is that it is not indicated whether 'power' in the definition is meant neutron power or thermal power. That is the reason why INSAG-7 claims that "[i]f neutron power is meant, then the personnel violated the Operating Procedures; if thermal power is meant, there was no violation" (INSAG-7 1992: 74).

Although it is questionable whether it can be considered as a factual violation of instruction of Operating Procedures, the action of raising the power after the drop played a serious role in the accident. The second report claims that "it is impossible to draw a final conclusion on whether or not the personnel actions were correct under these specific circumstances", but "the drop in reactor power at 00:28 and subsequent power increase were *largely to blame for the tragic consequences of the accident"* [italics are ours] (INSAG-7 1992: 74).

2. Responsibility, blameworthiness, and rule-violation in the INSAG-reports

The first INSAG report about the Chernobyl Unit 4 accident was seriously incomplete and biased against the operators due to the information presented to the IAEA by the Soviet authorities in 1986. The new information that came to light forced the agency to reconsider and restate its view. This information concerned the negative design features of Chernobyl-type reactors as well as design deficiencies of the control and safety rods. The second report had to compensate for the somewhat one-sided statements of the first report, especially concerning the evaluation of the rule-compliant behaviour of the operating personnel. The weight of responsibility given in INSAG-1 in 1986 "laid blame almost entirely on actions of the operating staff" (INSAG-7 1992: 24). By the time of the second report, this weight had to be lessened. So, the INSAG-7 report evaluates the formerly identified rule-violations of the operators one by one, in order "to shift the balance of perception so as to emphasize more the deficiencies in the safety features of the design which were touched on in INSAG-1, and to recognize the problems conferred by the framework within which plant operation was carried out. However, INSAG remains of the view that in many respects the actions of the operators were unsatisfactory" (INSAG-7 1992: 22).

The close reading of the contexts and the vocabularies of the re-evaluations of the formerly identified infringements of the operating personnel can reveal a kind of asymmetry. This asymmetry is related to the *blameworthiness* of the rule-breaking behaviour of the operating staff. When the second report based on the report of the USSR State Committee for the Supervision of Safety in Industry and Nuclear Power (SCSSINP) dismisses a charge of a concrete rule-breaking behaviour formerly identified in the first version, it typically uses the term "blame". For example, "the SCSSINP Commission does not support the apportionment of any blame to operating personnel" (INSAG-7 1992: 9), or "the Commission does not consider that personnel should be held to blame for having blocked the steam pressure protection system of the steam separators" (INSAG-7 1992: 12, 75-76), or "the Commission believes that the personnel cannot be blamed for disabling the reactor protection system which shuts down the reactor in the event of the closure of the emergency stop valves of both turbines" (INSAG-7 1992: 75), or "the Commission considers that the personnel cannot be held to blame for operating the unit at a power of less than 700 MW" (INSAG-7 1992: 76). So, the INSAG-7 nails down the cases when the operating personnel *cannot be blamed for* a rule-breaking action in the absence of an existing and relevant prescription or instruction. But, in contrast, the second report does not nail down those cases when the operating

personnel *can be* or *should be* blamed for a rule-breaking action. It seems that the INSAG-7 report avoids the term "blame" intentionally in relation to those cases when rule-breaking behaviour and actions were obvious, and they apply it to those cases only when the operators are exempted from a charge of an infringement.

However, INSAG-7 *implicitly* suggests that operators *can* or *should* be blamed in all cases where 'the actions of the operators were unsatisfactory', i.e., where there were existing and relevant instructions and the rule-breaking behaviors or actions were obvious. (INSAG-7 1992: 18).

3. Formal vs. Moral Responsibility

In general, the operators can be taken as blameworthy (or praiseworthy) for their actions for reasons from at least two different viewpoints.[5] First, they can be blame- or praiseworthy from a legal or formal viewpoint: whether they behaved in accordance with formal rules. Second, their behaviour or action can be evaluated from the viewpoint of moral responsibility: whether, on the ground that they acted freely and knew what they did, their actions were morally acceptable or not. Let us see these notions and their role in an explanation of the Chernobyl disaster.

3.1. Formal Responsibility

Formal or legal responsibility is derivable from a rule-following framework: in a rule-following environment, agents have a formal obligation to follow rules. Social activities require humans to follow rules that frame the activity in question. It goes without saying that a nuclear reactor is too complex to be analysed in a purely social framework. In rule-following contexts, the term 'social' is to be contrasted to 'individual' (Bloor 1997: 2-3, 84; Barnes 2000: 136) rather than 'physical' or 'technological'. Being social, such a framework is no less physical: social conventions are not matters of pure agreement in the Bloor-Barnesian context because the possible scope of the agreement is "limited by a system of severe existential coercion" (Fehér 2003: 178).

Actions in a nuclear reactor are essentially collective because, for obvious reasons, no one can individually operate a reactor. But rule-breaking is a *local* behaviour: individuals (or groups of

[5] Their epistemic responsibility can be a further aspect to be investigated (i.e., whether they have done their epistemic duties relevant for their actions), but we cannot do so that in the present paper, though we shall touch upon the epistemic *condition* of *moral* responsibility.

individuals) can break rules but a community as a whole cannot. If no one follows a rule, that rule is not constitutive of the game they play, though it may be constitutive of the game they are *supposed to* play, i.e., of the framework of (constantly broken) regulations. For example, in truth-telling games like science, *anyone* can cheat or lie, but from this, it does not follow that *everyone* can do so. On the contrary: if everyone lied, 'lie' would lose its meaning and 'liars' would start another game in which telling the truth is not a social requirement (Fehér 2003:175). This also relates to the *dilemma of Pilate* (Fehér 2003: 179): which is worse, Barabbas the breaker of rules, or Jesus the discreditor of rules? All the same, if a rule is broken by everyone, that is no rule for the game they follow, regardless of regulations external to their practice.

The other important aspect of this framework is the rules that set up the elbow room for agents. While humans are free to do what they wish to do within the framework of rules, if one breaks one of the (constitutive) rules of the activity, she gives up on that.[6] Some constitutive rules may be unbreakable due to the presence of the physical/technological environment. But as mentioned earlier (and Tanács and Danka (2020), as well as Danka and Tanács (2021), have shown to a greater extent), even some fundamental technological regulations were broken regularly on the ground that no direct consequences were identified prior to the disaster. A central question regarding the responsibility of operators in the case of Chernobyl is whether they were responsible for breaking the rules, or contributing to the disaster. As we shall see, the answer depends on a further question of whether formal or moral responsibility is at issue.

In the standard, Wittgensteinian version of rule-following, agents follow rules blindly – i.e., rule-following is non-reflective, and, so to say, unintelligently mechanic or automatic. As Wittgenstein famously put, "[w]hen I obey a rule, I do not choose. I obey the rule *blindly*" (Wittgenstein 1953/2009: 92). Even though there is often no rational explanation why rules must be followed, the only rational thing one can do in blind rule-following is following the rules blindly.

The reason is that (constitutive) rules are definitive about what is rational within the framework of

[6] Constitutive rules are to be distinguished from pure regulative rules. In the present context, the difference lies in the first being constitutive about the social framework of a practice to be followed, whereas the latter is not. In other words, breaking constitutive rules is giving up on the practice, whereas breaking regulative rules is less serious in consequences, being manageable within the original social framework. Not serving refreshement at breaks of a conference is breaking regulative rules of conferences, whereas not holding presentations at a conference is breaking constitutive rules.

practice and what is not: they are constitutive of the practice. Hence, rule-following is the ground rather than the consequence of rational behaviour: players do not follow the rules because they are rational, but rules are taken as the cornerstones of rational behaviour, and players' behaviour is taken as rational insofar as they follow the rules. As a consequence, questioning the rationality of a rule is possible only outside of the practice: e.g., within the framework of chess, it is not a legitimate move to ask why rules of chess are what they are, even if asking the same question may be reasonable outside of this framework (e.g., at chess class). As there is no ground for rationally questioning the rationality of rules within the game, there is no ground for their rejection either. Agents have two options: either they obey the rules, or they stop playing chess.

Blind rule-following does not necessarily imply being fully blind, non-reflective and unintelligent but it restricts the scope of reflection and intelligence to the elbow room within the framework of rules. Playing chess properly is a highly intelligent activity but its intelligent character does not occur in intelligently considering and evaluating the rules of chess. On the contrary: intelligence occurs within the framework of rules, i.e., the framework of acceptable moves within the game. Chess can be played intelligently: to be a clever chess player is to play chess properly, applying its rules, and making better moves than the opponent. Breaking the rules, in contrast, is an irrational action by definition: it means giving up playing the game that is, from the perspective of playing the game, the least rational thing to do.

To this framework, responsibility applies as a simple requirement: follow the rules blindly. This is how formal responsibility works: it is nothing else but compliance with rules. One is responsible for following the rules, and hence she is blameworthy if she misses to follow them.

A closely related interpretation of rule-following behaviour, as well as the responsibility associated with it, is commandant or superior responsibility. According to this interpretation, the violations of technological regulations discussed above would be merely questions of violating commands. From this point of view, a technological regulation is not different from a command, therefore following or infringing the former is equal to following or violating the latter. As a first step, the conception of the so-called commandant or superior responsibility assigns the duty as well as the responsibility to the superior "to properly supervise his subordinates" (Ronen 2010: 3). When a person follows a superior order in a way that she acts in accordance with the command then she is formally not responsible for the actions, including the consequences of her actions. But when she does not follow a superior order acting in accordance with the command or explicitly refuse the

command then she is formally responsible for violating the command.

There are some conditions when the principle of superior order does not release the command executor from liability. These are the conditions when the 'superior order defence' – i.e., claiming that someone is not responsible for her acts because she has "just followed orders" – is unacceptable, mainly because of the unacceptable consequences the agent must have foreseen. Although the rejection of this plea was formulated concerning war crimes, it can be generalised to a wider framework, in which a criminal or harmful act of a person under an order of a superior, whether military or civilian, shall not relieve that person of criminal or unlawful responsibility unless:

a) The person was under a legal obligation to obey orders of the superior in question;

b) The person did not know that the order was unlawful; and

c) The order was not manifestly unlawful (RSICC 1998: 19, Article 33; Sántha 2014: 317-318).

Conditions b) and c) mean that a person who executes a knowingly or manifestly unlawful superior command or follows a knowingly or manifestly illegal regulation is responsible for the criminal or harmful consequences of her rule-following behaviour. But when a person does not have a knowledge of the illegality of an order/regulation and the order/regulation was not manifestly illegal then she cannot be responsible for her rule-following behaviour.

Both the Wittgensteinian description and the command-based conception capture the (almost) unconditional character of rule-following. According to that, all that an agent needs to know is what to do under the relevant circumstances and how to do so. No deliberation, no consideration, but just taking the unconditional order: under circumstances *c*, always do action *a*. This kind of conception basically expects or supposes obedience-based rule-following behaviour. Since obedience does not require any specific intellectual insight why to follow a particular command or regulation, the wording of a command or regulation typically does not offer any reason or explanation.

A central claim of ours is that a command-based approach is insufficient for proper rule-following in the Chernobyl case. At least three objections can be raised against the applicability of that approach to the case of the Chernobyl disaster. These objections, strengthening each other, seem for us to imply that the above-discussed primitive conception of rule-following is not suitable for

explaining rule-breaking behaviour in the Chernobyl case, and it is also not suitable for providing an evaluative framework for the moral responsibility of the operators. Here we discuss the first of our three objections as they can be understood on the ground of the above, whereas to put the other two into context, we need to introduce the notion of moral responsibility in a bit more detail. This shall be done in the next section.

Objection 1. Expert opinion suggests that command-based rule-following is insufficient for controlling, leading, and managing a civilian- (i.e., non-military-)operated nuclear reactor. Szatmáry and Aszódi claim that nuclear energy production has evolved typically as a result of military application all over the world, and in the case of the Soviet Union, the entire industry remained under military control for much longer than in other countries (Szathmáry and Aszódi 2005: 95).

Since the RBMK-type nuclear reactor as of Chernobyl was advantageous for the production of plutonium for military use, the reactor could have been used for the production of both electricity and plutonium. The problem is that civilian-need energy production and military-need plutonium production require a sharply different mode of operation. Even though the reactors of Chernobyl were operated specifically for civilian purposes prior to the accident, switching the mode to military use could not have been a big problem. All this enhanced the role of the military mindset.

Due to the origin of military application and military thinking pervading the operation, it was supposed that the coercive force of the operating procedures is high enough and sufficient for avoiding rule-breaking. But it turned out that the coercive force of the regulations in civilian areas was much lower than in military areas. So, as "the example of Chernobyl shows that civilian operations personnel can be persuaded by certain factors to violate the regulations" (translations are ours, Szathmáry and Aszódi 2005: 95).[7]

As further evidence supporting this objection, INSAG-7 considered the surviving military approach to be the source of a problematic and mistaken attitude. The report of INSAG quoted a note from the original soviet material. In the document, the representatives of the Scientific Manager identified the mistakenness of conception in the following way:

Many years of accident free operation of military reactors in the USSR led to the deep rooted

[7] In relation to the detailed analysis of that what these factors were and what role they could play in the infringements, see Danka-Tanács 2021.

philosophy that one only has to write correct reactor operating instructions to guarantee safety. It goes without saying that these instructions are mandatory for the personnel. It turned out that the real situation was very different. The first most important lesson to be learnt from Chernobyl is that the safety of nuclear power plants cannot be based on instructions (INSAG-7 1992: 82).

The IAEA seemed to accept the assessment of the situation but considered adding a correction suggesting that "it is difficult to claim that the RBMK operating instructions in force prior to 1986 were correct" (INSAG-7 1992: 82). In the light of this correction, it can be said that the mandatory characteristic of the instruction is necessary but not sufficient, and in the case of Chernobyl, the correctness of the instructions can be questioned and reviewed.

The expert opinion cited above suggests that the conception of the command-based rule-following is not adequate and hence cannot be accepted from a pragmatical viewpoint when rule-following behaviour is expected in a non-military area. We also have two philosophical problems with this account. The introduction of them, as mentioned above, requires some background understanding of moral responsibility and its conditions.

3.2. Entering the Sphere of Morality: The Control Condition

Moral responsibility is taken to have two conditions: a so-called control condition (or freedom condition) and an epistemic condition (more precisely, a set of four epistemic conditions). The control condition requires agents to have control over their acts, and the epistemic condition requires them to be aware of the action, its moral significance, possible consequences and possible alternatives. Without satisfying these conditions, an agent cannot be responsible for her acts (Talbert 2019).

The Chernobyl case is very instructive about the control condition. Blind, command-based rule-following situations normally suppose a deterministic environment: rules determine the elbow room of agents. Even if they are physically capable of breaking rules, this also involves a breakup with the social framework establishing those rules. This is why e.g. breaking the law can have consequences to one's social status.[8]

[8] One may argue that in the case of Chernobyl, a much more serious consequence than losing social status is the possibility of a nuclear disaster. However, Soviet propaganda made even nuclear engineers believe that their reactors are super-safe and a disaster is literally impossible (Danka and Tanács 2021).

But from the very moment of their rule-breaking, the operators made operationally distinguishable choices: they broke operational regulations and hence in this sense, they gained control over their actions. More precisely: by breaking the rules, they broke out of the operationally deterministic environment they were supposed to be involved in. By this, they made themselves subject to moral judgment (on the control condition side at least – see again Talbert 2019). In contrast, if they had been committed to rule-following behaviour, they could not have been taken responsible for their acts in an operational sense of the control condition. They had orders from their leaders, they had instructions to follow, they were ordered to do a test on the reactor in order to make it certain that the reactor operates accident-free, etc.

In this operational sense of control and freedom, the operators can be taken to be responsible for their actions in the cases and only in the cases when they violated operational rules (if the epistemic condition were also met which were not, as we shall see in Sect. 4). They cannot be taken responsible for following possibly inappropriate orders that may lead to potentially blameworthy consequences since by committing themselves to rule-following behaviour, they had no choices (with the above-mentioned exceptions that do not apply to the Chernobyl case). The choice they made was logically prior to *doing* this or that: it was the choice between rule-following and rule-ignorance.

Following instructions in a nuclear reactor is mainly a technological-operational (rather than moral) matter. Following rules 'blindly' is normally required for operational reasons like in the case of e.g. military or surgery operations. On the moral side, an important consequence of blind rule-following is precisely taking responsibility off from the shoulders of rule-followers. Breaking rules is therefore not only significant formally as it makes rule-breakers formally blameworthy, but it is also morally significant because, by rule-breaking, rule-breakers take responsibility for their actions, even though they could hide behind regulations that would make them responsible neither in the formal nor the moral sense. Examples of rule-breaking in Chernobyl, as we shall see in the following section, included cases when, in our understanding, rule-breaking was intentional and taking responsibility was deliberate. But even in these cases, the moral responsibility of the operators is, as we shall also see, questionable based on the epistemic condition (Sect. 4.).

Objection 2. The control condition of moral responsibility, as applied to a blind rule-following framework, implies that agents cannot be taken to be morally responsible as far as they follow the regulations. But once they had given up on their commitment to rule-following, they had

alternatives to choose from. In cases when they follow orders, the responsibility is of those who had given the orders. Breaking a rule implies that they free themselves from (some of the) regulations, making the control condition satisfied. In the present context, control and freedom are relative to the operational regulations of the power plant:[9] an operator is in control of her acts if and only if she breaks some regulations and hence decides to gain authority over her acts relevant to the rule-breaking. This condition of moral responsibility was met in the case of Chernobyl, but for the price of giving up blind rule-following, i.e., the framework that was supposed to govern all activities in the Chernobyl reactor on the one hand, and was also held by a philosophical account of rule-following that is inapplicable to Chernobyl-like cases on the other hand.

4. Moral Responsibility and the Epistemic Condition

Meeting the control condition is necessary but insufficient for moral responsibility. It opens up the door to moral evaluation but as it has been mentioned earlier, epistemic condition(s) must be also satisfied. More precisely, four components of the epistemic condition, incl. an awareness of (1) action, (2) consequences, (3) alternatives, and (4) moral significance must be met. In rule-following situations, following rules consciously requires a knowledge and understanding of these components. As the notion of the epistemic condition of moral responsibility has been introduced, the third objection mentioned above can be also discussed.

Objection 3. Requirements of awareness of various factors cannot be met in a blind rule-following framework, the essence of which is precisely blindness to these components of the epistemic condition for moral responsibility. Whether epistemic condition(s) of responsibility were satisfied in the case of rule-violations before the Chernobyl disaster is a question that cannot be discussed here in sufficient depth. But one thing is certain: regulations were not clear about some of the components of the epistemic condition, incl. possible consequences of rule-breaking. A reason might be that for blind rule-following, no epistemic condition must be satisfied other than knowledge and understanding of rules, those of the action expected, and the action's being fit into the framework of rules. If no rule violation ever happens because everyone follows the rules blindly, there is no need to know the consequences, alternatives, and moral significance of the planned actions. But since blind rule-following cannot be warranted by command-based rules

[9] This also implies that our analysis has nothing to do with problems of metaphysical notions of freedom and control.

alone, it is indispensable for rule-makers to formulate rules so that other components (and, first of all, possible consequences of rule-breaking) are also included in the regulations. Being unaware of the consequences of their actions in advance, and reinforced by previous experience that their actions have no negative consequences, operators cannot be taken as morally responsible for the disaster.

5. Conclusion

We have argued that a blind, command-based rule-following framework is not suitable for explaining rule-breaking behaviour in Chernobyl, nor is it suitable for evaluating the moral responsibility of rule-breakers. Drawing the consequences from INSAG-7 as the historical resource widely held to be the most accurate as well as the expert opinion still taken to be valid today, we have come to the conclusion that operators are to be taken as *formally* responsible for their rule violations, because they violated regulations that contributed to the disaster. However, the responsibility attribution should be radically decreased in comparison with the evaluation of earlier reports because external factors (most notably, regulation issues) also played their role in the accident and rule violations were certainly not root causes. INSAG-7 and other historical resources refrain from a *moral* evaluation of the operators' actions. This task has been up to us. We have found that even if the operators' wrongdoings served a ground for formal responsibility of theirs, and even if intuitions may imply that slightest rule violations leading to a disaster of that calibre should serve as a ground for moral responsibility, conditions of moral responsibility widely accepted by theories of responsibility were not satisfied. Hence, the operators cannot be taken as morally responsible for the consequences of their rule violations, though they can be taken as formally responsible for the violations themselves.

We have raised three objections against the applicability of the blind rule-following framework to the case of Chernobyl. First, blind rule-following is designed mainly for military purposes, and its applicability to non-military contexts is questionable at least. Second, in such a framework, commands rather than the agents themselves control acts, and hence the control condition of moral responsibility cannot be satisfied. Giving up the blind rule-following framework, the control condition can be satisfied, as it happened in the case of the disaster. However, as the third objection demonstrates, operators did not have sufficient information about the possible

consequences of their possible rule-breaking, and hence (one of) the epistemic condition(s) of moral responsibility was not met. A reason for this underinformedness seems to us that operators were expected to follow blind, command-based rules, and a possibility of rule-breaking was not even taken into account. However, this expectation was proved to be groundless on the one hand, and also insufficient for managing emergency situations in the case of eventual rule-breakings on the other.

As we have also shown, the epistemic conditions of moral responsibility imply certain expectations regarding the formulation of regulations. Disregarding epistemic conditions of moral responsibility by an inappropriate wording of rules, regulators had released operators' of responsibility in advance: since operators did not know some epistemically relevant factors for their acts like the consequences of breaking certain rules, they could not be taken as morally responsible for those consequences (even though they are to be taken as formally responsible for the lesser misdeed, namely breaking rules).

References

Barnes, Barry. 2000. Understanding Agency. Social Theory and Responsible Action. London: Sage.

Bloor, David. 1997. Wittgenstein, Rules and Institutions. London: Routledge.

Coren, Daniel. 2019. Freedom, Gratitude, And Resentment: Olivi And Strawson. *Res Philosophica*. 96 (3): 307–327.

Danka, István and Tanács, János. 2021. Rationalising Rule Violation in the Case of Chernobyl Disaster: Six Systematic Excuses. *Információs Társadalom*. 21 (4): 19-37.

Frankfurt, Harry. 1969. Alternate Possibilities And Moral Responsibility. *Journal of Philosophy*. 66 (23): 829–839.

Fehér, Márta. 2003. Az episztemológia morális dimenziója (Polányi Mihály

tudományfilozófiájáról). In Farkas, K. Orthmayr, I. Bölcselet és analízis. Budapest: ELTE Eötvös Kiadó. 165-181.

INSAG-7. 1992. *The Chernobyl Accident: Updating of INSAG-1. A Report by the International Nuclear Safety Advisory Group.* Safety Series No. 75. Vienna: International Atomic Energy Agency. https://www-pub.iaea.org/MTCD/publications/PDF/Pub913e_web.pdf

Margitay, Tihamér and Ziegler, Zsolt. 2014. Alternative Possibilities, Self-Determination And Responsibility, *Polanyiana* 23 (1-2): 40-50. http://www.polanyi.bme.hu/folyoirat/2014-1/Margitay-Ziegler.pdf

Medvedev, Zhores A. 1990. *The Legacy of Chernobyl*. W. W. Norton & Company, Inc.: New York.

Ronen, Yael. 2010. Superior Responsibility of Civilians for International Crimes Committed in Civilian Settings. *Vanderbilt Journal of Transnational Law.* 43 (2): 1-38.

RSICC. *Rome Statute of the International Criminal Court*. 1998. https://legal.un.org/icc/statute/english/rome_statute(e).pdf

Rudy-Hiller, Fernando. 2019. The Epistemic Condition for Moral Responsibility. In Zalta 2019. https://plato.stanford.edu/archives/win2019/entries/moral-responsibility-epistemic/

Sántha, Ferenc. 2014. Az elöljáró parancsa mint büntethetőségi akadály a magyar büntetőjogban. *Publicationes Universitatis Miskolcinensis. Sectio Juridica et Politica.* 32: 313–320. http://midra.uni-miskolc.hu/document/19874/13860.pdf

Szathmáry, Zoltán and Aszódi, Attila. 2005. *Csernobil. Tények, okok, hiedelmek*. Budapest: Typotex Kiadó.

Tanács, János and Danka, István. 2020. Mulasztás és engedetlenség: a csernobili atomkatasztrófa néhány tanulsága. Gradus 7: 131-140.

Talbert, Matthew. 2019. Moral Responsibility. *In:* Zalta 2019. https://plato.stanford.edu/archives/win2019/entries/moral-responsibility/

Wittgenstein, Ludwig. 1953/2009. *Philosophische Untersuchungen. Philosophical Investigations*. Wiley-Blackwell.

Zagzebski, Linda and Fairweather, Abrol (eds.) 2001. *Virtue Epistemology.* Oxford: Oxford University Press.

Zalta, Edward N. 2019. *The Stanford Encyclopedia of Philosophy.* https://plato.stanford.edu/

Ziegler, Zsolt. 2016. A Relational Theory Of Moral Responsibility. *Prolegomena: Journal of Philosophy* 15 (1): 71-88.