1. **Background: strong tendencies towards scepticism**

Notion of retrospective responsibility closely tied to notion of desert: we deserve rewards and punishments or blame and praise for something only if we are responsible for it. Central moral and social notion.

Sceptical arguments: Galen Strawson, Peter van Inwagen, Derk Pereboom, Al Mele, etc

**Tendencies towards scepticism:** We are prone to find moral responsibility incompatible with both determinism and indeterminism, making metaphysical demands going well beyond our explicit concerns in everyday attribution of moral responsibility.

Standard defences: undermine arguments.

But: tendency to scepticism seems independent of particular arguments. Direct response to deterministic scenarios.

Strategy: “explain away” both general tendency and reaction to particular arguments.

2. **Conceptual mechanics and aetiology: The Explanation Hypothesis**

Practices of holding responsible have purpose: (i) Explicit justification, (ii) Resentment sensitive to whether object is of a kind that can change in response, (iii) Guilt has change-inducing psychological function, (iv) Resentment placated by change or willingness to change, (v) HR directed at third parties aims at confirmation of corresponding attitudes.

--- diagram over the etiology of responsibility judgments ---

- concept of moral responsibility
- produces judgments of moral responsibility
- guide practices of holding responsible
- have as their function to support
- General Explanatory Tendency (GET)
- Reactive Response-ability (RR)
- Explanatory Responsibility (ER)
- which requires
- normative expectations
- appropriate motivational structures

**EH:** We attribute moral responsibility insofar as we take GET, RR and ER to be satisfied

---

**The Explanation Hypothesis (EH):** People take P to be morally responsible for E to the extent that they take E to be an event of a type T and take P to have a motivational structure S of type M such that GET, RR, ER and MV hold:

**General Explanatory Tendency (GET):** Type M motivational structures are part of a reasonably common sort of significant explanation of type T outcomes.

**Reactive Response-ability (RR):** Type M motivational structures tend to result from or respond in the right way to agents being held responsible for realizing or not preventing type T events.

**Explanatory Responsibility (ER):** S is part of a significant explanation of E of the sort mentioned in GET.

**Moral Valance (MV):** Type T events have moral valance.

- E: intentions, decisions, actions, causal consequences of actions of things we fail to prevent.
- MV captures moral responsibility. EH without MV characterizes “agential responsibility”.


What EH accounts for:

**A.** People are seen as morally responsible for most sought for morally good or bad outcomes.

**B.** Compulsion, phobias, severe personality disorders and reactions to extreme stress are seen as responsibility undermining.

**C.** Degrees of external force, threats and costs are seen as variously mitigating moral responsibility.

**D.** Ignorance is seen as mitigating moral responsibility, but not always.

**E.** Those who actively bring about an outcome are more responsibility for it than those who allow others to bring it about.

**F.** People are seen as collectively responsible for outcomes over which they had no individual control.

**G.** Strong correlation between judgements of MR and certain explanatory judgments.

**H.** Judgments of moral responsibility tend to be undermined by reductionistic, mechanistic explanations of behaviour.

**I.** Most people ascribe a higher degree of responsibility for negative than positive side effects that the agent did not care about.
3. Explaining sceptical intuitions

Explanatory frames affect what is taken as a significant explanation:

(1) Relevant explanatory models: level of abstraction or detail, kind of explanation.
(2) Various practical or theoretical interests.
(3) Contrastive concerns.

Applications:

J. The intuition that causal determinism undermines moral responsibility. (i) Abstracts away from details. (ii) Focuses on prior causes. Focus on prior cause somewhat effective in itself in changing explanatory frame; abstraction removes one constraint on this.

K. The appeal of regress arguments (G. Strawson, van Inwagen). Similar mechanisms as in J.

L. Arguments from manipulation. The motivational structures of manipulators are often significant explanations; these arguments invite us to see other causes of agents’ motivation and action as similarly significant.

M. Contrastive arguments from luck. Contrastive arguments invite us place the relevant motivational structures in the explanatory background.

All these sceptical arguments induces explanatory frames different from those of everyday social interaction from the participatory stance.

4. The illusion of undermined responsibility

Even though the horizontal lines are parallel, they do not seem to be:

![Parallel lines](image)

Parallel: Under some conditions, agents are, but do not seem to be morally responsible.

EH provides ground for attributing error one way rather than another. The mechanism explaining sceptical intuitions is one of dispelling without disqualifying.

(1) Given frame relativity, there is no univocal analytic concept of moral responsibility.

(2) Second, the aetiological story that motivates EH and is supported by its success suggests that our concept or moral responsibility nevertheless has a definite cognitive function: to identify when the right sort of motivational structures stand in the right explanatory relation to events.

(3) Third, in order to reinforce relevant normative expectations, practices of holding responsible needs to be directed at cases where the explanatory relation between motivational structure and object of responsibility is straightforward relative to stable ordinary, everyday perspectives.

Given the evidence for EH, then, our best understanding of our concept of moral responsibility do not take it to be hostage to standard sceptical worries. When it seems that it does, this is best understood as yet another example of the human mind’s susceptibility to illusions.

5. Objections

(1) Everyday perspective one of ignorance; sceptical perspective approaches God’s eye view.

   a. Sceptical perspectives equally selective, just differently.
   b. Selectivity and ignorance are not disqualifiers.

(2) Some people have more definite concepts, incorporating a variety of further constraints, and perhaps constraints that require uncaused causes or agent causation.

   a. Their concept might not be distinct. Deferential intentions.
   b. The real question becomes one of determining when and how to hold people responsible. Not clear that their sceptical intuitions can play an independent role.

6. Related works

Björnsson, G 2011 “Joint responsibility without individual control: Applying the Explanation Hypothesis”, forthcoming in Compatibilist Responsibility: beyond free will and determinism, eds. Jeroen van den Hoven , Ibo van de Poel and Nicole Vincent
