

Peter Geach: Das Frege-Geach Problem

Kaj André Zeller

University of Leeds

2025-12-04

In einer Reihe von Aufsätzen entwickelten Peter Geach (1958, 1960, 1965) und John Searle (1962) einen bedeutenden Einwand gegen nichtkognitivistische metaethische Positionen. Zu jener Zeit war der Nichtkognitivismus vorherrschend, und der von Geach und Searle formulierte Einwand trug entscheidend dazu bei, diesen zu hinterfragen. Das Problem, wie der Nichtkognitivismus mit diesem Einwand umgehen kann, ist heute als das ‚Frege-Geach-Problem‘ bekannt, da Geach den Einwand auf Gottlob Freges (1879) Unterscheidung zwischen Inhalt und Urteil zurückführte. Das Frege-Geach-Problem hat sowohl die Verteidiger nichtkognitivistischer Positionen als auch deren Kritiker intensiv beschäftigt und wird häufig als zentrales Problem in der Entwicklung neuerer nichtkognitivistischer Theorien angesehen. Im Folgenden rekonstruiere ich die bekannteste Version dieses Einwands, die sich in erster Linie gegen die Theorie richtet, dass es nicht beschreibend, sondern verurteilend ist, eine bestimmte Handlung als ‚schlecht‘ zu bezeichnen. Diese Theorie wird von Geach R.M. Hare (1952) zugeschrieben.

Kaj André Zeller: “Peter Geach: Das Frege-Geach Problem”; *argumentation.online*, 2025-12-04, www.argumentation.online/pdfs/Zeller_ArgOnl-2025-12.pdf. Veröffentlicht unter der Creative Commons Lizenz (by-nc).

Bibliographische Angaben

Geach, P. (1965). Assertion. *Philosophical Review*, 74(4), 449–465. <https://phil-papers.org/archive/GEAA-2.pdf>

Textstelle

»... an ostensibly assertoric utterance “p” and “If p, then q” can be teamed up as premises for a modus ponens. Here, the two occurrences of “p,” by itself and in the “if” clause, must have the same sense if the modus ponens is not to be vitiated by equivocation; and if any theorist alleges that at its ostensibly assertoric occurrence “p” is really no proposition at all, it is up to him to give an account of the role of “p” that will allow of its standing as a premise ... The theory that to call a kind of act “bad” is not to describe but to condemn it is open to similar objections. Let us consider this piece of moral reasoning: If doing a thing is bad, getting your little brother to do it is bad. Tormenting the cat is bad. Ergo, getting your little brother to torment the cat is bad. The whole nerve of the reasoning is that “bad” should mean exactly the same at all four occurrences—should not, for example, shift from an evaluative to a descriptive or conventional or inverted-commas use. But in the major premise the speaker (a father, let us suppose) is certainly not uttering acts of condemnation: one could hardly take him to be condemning just doing a thing.« (463-4)

»... eine scheinbar assertorische Äußerung „p“ und die Aussage „Wenn p, dann q“ können zusammen als Prämissen für einen Modus Ponens dienen. Dabei müssen die beiden Vorkommnisse von „p“, einmal allein und einmal im „wenn“-Satz, dieselbe Bedeutung haben, damit der Modus Ponens nicht durch Mehrdeutigkeit ungültig wird; und falls irgendein Denker behauptet, dass bei seiner scheinbar assertorischen Verwendung „p“ gar keine Proposition sei, liegt es an ihm, eine Erklärung für die Rolle von „p“ zu liefern, die es erlaubt, es als Prämissen zu betrachten ... Die Theorie, dass es nicht beschreibend, sondern verurteilend ist, eine bestimmte Handlung als „schlecht“ zu bezeichnen, ist ähnlichen Einwänden ausgesetzt. Betrachten wir dieses moralische Argument: Wenn es schlecht ist, eine Handlung auszuführen, dann ist es schlecht, deinen kleinen Bruder dazu zu bringen, diese Handlung auszuführen. Es ist schlecht, die Katze zu quälen. Also ist es schlecht, deinen kleinen Bruder dazu zu

bringen, die Katze zu quälen Der Kern der Argumentation ist, dass „schlecht“ an allen vier Stellen genau dieselbe Bedeutung hat—es sollte zum Beispiel nicht von einer wertenden zu einer beschreibenden, konventionellen oder ironischen Verwendung wechseln. Aber in der Hauptprämisse [welche da ist: „Wenn es schlecht ist, eine Handlung auszuführen, dann ist es schlecht, deinen kleinen Bruder dazu zu bringen, diese Handlung auszuführen“; Anmerkung des Verfassers] äußert der Sprecher (nehmen wir an, ein Vater) sicherlich keine Verurteilung: Man könnte kaum annehmen, dass er damit einfach nur das Ausführen einer Handlung verurteilt.« (Übersetzung des Verfassers)

Argumentrekonstruktion

1. Das folgende Argument ist gültig: Hauptprämisse: Wenn es schlecht ist, eine Handlung auszuführen, dann ist es schlecht, deinen kleinen Bruder dazu zu bringen, diese Handlung auszuführen. Nebenprämisse: Es ist schlecht, die Katze zu quälen. Konklusion: Also ist es schlecht, deinen kleinen Bruder dazu zu bringen, die Katze zu quälen.
2. Wenn dieses Argument gültig ist, hat „schlecht“ dieselbe Bedeutung in der Hauptprämisse wie in der Nebenprämisse.
3. In der Hauptprämisse liegt die Bedeutung von „schlecht“ nicht darin, eine Verurteilung auszudrücken.
4. Nach der Theorie, dass es nicht beschreibend, sondern verurteilend ist, eine bestimmte Handlung als „schlecht“ zu bezeichnen, liegt die Bedeutung von „schlecht“ in der Nebenprämisse darin, das Quälen der Katze zu verurteilen.

5. Die Theorie, dass es nicht beschreibend, sondern verurteilend ist, eine bestimmte Handlung als „schlecht“ zu bezeichnen, ist falsch.

Kommentar

Die obige Rekonstruktion von Geachs Argumentation ist gültig, wie in der untenstehenden Detailanalyse zu sehen ist. Das bedeutet, dass die Wahrheit der Prämissen die Wahrheit der Konklusion nach sich zieht. Ob das Argument auch schlüssig bzw. stichhaltig ist, hängt davon ab, ob die Prämissen tatsächlich

wahr sind bzw. ob wir gute Gründe haben anzunehmen, dass die Prämissen wahr sind.

Prämissen 1 und 2 in dieser Rekonstruktion von Geachs Argument sind sehr plausibel, auch wenn ein Nichtkognitivist natürlich behaupten könnte, dass moralische Ausdrücke sich einfach nicht dafür eignen, deduktive Schlüsse zu ziehen. Prämissen 4 folgt aus der Definition der Theorie, die Geach angreift. Aber ist Geachs Darstellung dieser Theorie, die er Hare zuschreibt, auch gerecht? Und inwiefern lässt sich das Argument auf nichtkognitivistische Theorien im Allgemeinen übertragen?

Die wichtigste Prämissen in dieser Rekonstruktion von Geachs Argument scheint allerdings zunächst Prämissen 3 zu sein. Das Problem, das sich für den Nichtkognitivismus daraus ergibt, einen Weg zu finden, Prämissen 3 abzulehnen, lässt sich wie folgt ausdrücken: Nichtkognitivisten müssen darlegen, wie moralische Ausdrücke in allen Verwendungskontexten oder „Einbettungen“ dieselbe Bedeutung beibehalten können. Aus diesem Grund ist das Frege-Geach-Problem auch als „Einbettungsproblem“ (embedding problem) bekannt. Das Hauptproblem, das Geachs Argument aufzuzeigen scheint, liegt darin, dass, wenn der Ausdruck „schlecht“ in einen „wenn“-Satz eingebettet wird, er nicht die Bedeutung anzunehmen scheint, von der der Nichtkognitivismus behauptet, dass sie in einem scheinbar assertorischen Satz vorliegt. „Wenn“-Sätze sind jedoch nur ein Beispiel für Einbettungen. Der Nichtkognitivismus sollte zum Beispiel auch die Bedeutung moralischer Ausdrücke in Fragen (zum Beispiel: „Ist es schlecht, die Katze zu quälen?“), in Wahrheitszuschreibungen (zum Beispiel: „Es ist wahr, dass es schlecht ist, die Katze zu quälen“) und in Negationen (zum Beispiel: „Es ist nicht der Fall, dass es schlecht ist, die Katze zu quälen“) erklären können.

Selbstverständlich gibt es verschiedene nichtkognitivistische Ansätze, das Frege-Geach-Problem zu lösen. Ein Ansatz besteht darin, eine „Logik mentaler Einstellungen“ zu entwickeln. Dieser Ansatz beruht auf der Beobachtung, dass nichtkognitive Einstellungen, wie zum Beispiel Wünsche, Verlangen und Begehren, in dissonanten Verhältnissen stehen können.

Eine Version dieses Ansatzes besteht darin, sich auf höherstufige mentale Einstellungen zu beziehen. Ausgehend von Blackburns (1984) Überlegungen lässt sich folgender Lösungsansatz entwickeln: Nichtkognitivisten können

zunächst Prämissen 4 des obigen Arguments akzeptieren: Dem Nichtkognitivismus zufolge drückt der Satz „Es ist schlecht, die Katze zu quälen“ tatsächlich eine Verurteilung des Katzenquälens aus. Prämissen 3 jedoch müssen Nichtkognitivisten nicht zustimmen. Blackburn schlägt vor, die Bedeutung von Konditionalen wie folgt zu erklären: „wenn“-Sätze drücken höherstufige, nichtkognitive Einstellungen aus, die sich auf andere mentale Einstellungen beziehen. Der Satz „Wenn es schlecht ist, eine Handlung auszuführen, dann ist es schlecht, deinen kleinen Bruder dazu zu bringen, diese Handlung auszuführen“ drückt beispielsweise die Verurteilung einer bestimmten Kombination von Einstellungen aus—nämlich der Kombination, eine Handlung zu verurteilen, aber nicht zu verurteilen, den kleinen Bruder dazu zu bringen, diese Handlung auszuführen. Ein Nichtkognitivist könnte so behaupten, dass in diesem Sinne die Bedeutung von „schlecht“ auch in der Hauptprämissen darin liegt, eine Verurteilung auszudrücken.

Selbstverständlich wurden auch Einwände gegen nichtkognitivistische Lösungsansätze erhoben. Einflussreich ist beispielsweise Van Roojens (1996) Einwand gegen eine „Logik mentaler Einstellungen“. Van Roojen argumentiert, dass Blackburns Lösung logische und praktische Widersprüchlichkeit zu vermengen scheint. Der Satz „Es ist schlecht, eine Handlung zu verurteilen und nicht zu verurteilen, deinen kleinen Bruder dazu zu bringen, diese Handlung auszuführen“ drückt die Verurteilung der mentalen Einstellung aus, gleichzeitig eine Handlung zu verurteilen, aber nicht zu verurteilen, deinen kleinen Bruder dazu zu bringen, diese Handlung auszuführen. Blackburn zufolge drückt das Konditional „Wenn es schlecht ist, eine Handlung auszuführen, dann ist es schlecht, deinen kleinen Bruder dazu zu bringen, diese Handlung auszuführen“ genau dasselbe aus.

Demzufolge müsste das folgende Argument gültig sein:

Hauptprämissen: Es ist schlecht, eine Handlung zu verurteilen und nicht zu verurteilen, deinen kleinen Bruder dazu zu bringen, diese Handlung auszuführen. Nebenprämissen: Es ist schlecht, die Katze zu quälen. Konklusion: Also ist es schlecht, deinen kleinen Bruder dazu zu bringen, die Katze zu quälen.

Allerdings ist dieses Argument nicht gültig. Insofern ergibt sich ein Problem zumindest für Blackburns oben vorgestellten Lösungsansatz.

Ein weiterer, zunächst von Unwin (1999) entwickelter Einwand gegen Blackburns Lösungsansatz richtet die Aufmerksamkeit auf Negationseinbettungen. Da sich dieser Einwand als besonders problematisch erwiesen hat, ist er unter dem Namen „Negationsproblem“ bekannt geworden. In der Tat wird das Frege-Geach-Problem aktuell häufiger im Zusammenhang mit Negationen als mit Konditionalen diskutiert.

Ausgehend von Unwins Überlegungen können wir feststellen, dass ein Satz wie „Ich glaube, dass es falsch ist, die Katze zu quälen“ drei Möglichkeiten einer Negationseinbettung zulässt:

- (i) Ich glaube nicht, dass es falsch ist, die Katze zu quälen.
- (ii) Ich glaube, dass es nicht falsch ist, die Katze zu quälen.
- (iii) Ich glaube, dass es falsch ist, die Katze nicht zu quälen.

Das Negationsproblem ergibt sich nun daraus, dass der Satz „Ich verurteile, die Katze zu quälen“ nur zwei Negationseinbettungsmöglichkeiten zulässt:

- (i) Ich verurteile nicht, die Katze zu quälen.
- (ii) ???
- (iii) Ich verurteile, die Katze nicht zu quälen.

Schroeder (2008) schlägt vor, das Negationsproblem wie folgt anzugehen. Zunächst nimmt er an, dass es eine nichtkognitive mentale Einstellung gibt, welche die folgende Eigenschaft besitzt: Wenn der Inhalt von zwei Instanzen dieser Einstellung inkonsistent ist, dann stehen die beiden Instanzen dieser Einstellung selbst im Widerspruch zueinander. Schroeder nennt diese Einstellung stipulativ „Dafür-sein“ („Being-for“). Nach Schroeder lässt sich das Negationsproblem dann folgendermaßen lösen:

- (i) Ich bin nicht dafür, die Katze zu quälen.
- (ii) Ich bin dafür, die Katze nicht zu quälen.
- (iii) Ich bin dafür, zu verurteilen, die Katze zu quälen.

Mithilfe dieses Ansatzes lässt sich der Widerspruch zwischen den folgenden Sätzen wie folgt als Widerspruch zwischen mentalen Einstellungen denken:

- „Es ist schlecht, die Katze zu quälen“ / „Es ist nicht der Fall, dass es schlecht ist, die Katze zu quälen“ = Dafür-sein, das Quälen der Katze zu verurteilen /

Nicht Dafür-sein, das Quälen der Katze zu verurteilen • „Es ist schlecht, die Katze zu quälen“ / „Es ist nicht schlecht, die Katze zu quälen“ = Dafür-sein, das Quälen der Katze zu verurteilen / Dafür-sein, das Quälen der Katze nicht zu verurteilen

Auf dieser Basis entwickelt Schroeder ein komplexes System, um weitere Einbettungen, wie zum Beispiel Einbettungen in „wenn“-Sätze und Einbettungen in „es ist wahr, dass“-Sätze, zu erklären. Es ist nun aber zu erwähnen, dass Schroeders „Dafür-sein“-Variante des Nichtkognitivismus eine Weiterentwicklung älterer „Verurteilung“-Varianten des Nichtkognitivismus ist. Nach Schroeders Variante drückt „Es ist schlecht, die Katze zu quälen“ keine Verurteilung, sondern ein Für-eine-Verurteilung-sein aus. Demzufolge beruht Schroeders Nichtkognitivismus nicht unmittelbar auf der Theorie, wie Geach sie beschreibt, und ist daher nicht direkt von Geachs Kritik betroffen.

Neben der „Logik mentaler Einstellungen“ gibt es einen weiteren einflussreichen Ansatz, das Frege-Geach-Problem zu lösen. Dieser Ansatz besteht darin, eine deflationäre Wahrheitstheorie in Anspruch zu nehmen – das heißt, eine Wahrheitstheorie anzunehmen, nach der ein Satz „p“ genau dann wahr ist, wenn p. Zum Beispiel: „Schnee ist weiß“ ist genau dann wahr, wenn Schnee weiß ist.

Stoljar (1993) schlägt beispielsweise den folgenden Lösungsansatz für den Emotivismus, eine Unterart des Nichtkognitivismus, vor. Seiner Ansicht nach sollten Emotivisten scharf zwischen zwei Thesen ihrer Position unterscheiden, nämlich zwischen der semantischen These, dass moralische Sätze nur in einem deflationären Sinne wahr sind, und der pragmatischen These, dass moralische Aussagesätze nicht die Funktion haben, die Welt zu beschreiben, sondern Emotionen auszudrücken. Mithilfe dieser Unterscheidung kann der Emotivist argumentieren, dass es zwar zutrifft, dass nach dem Emotivismus die Haupt- und Nebenprämissen des obigen Arguments nicht dieselbe emotionale Bedeutung haben, dass aber dieser Bedeutungswechsel keine Auswirkung auf die Gültigkeit des Arguments hat; er ist lediglich pragmatisch, nicht semantisch. Die Gültigkeit des Arguments, so Stoljar, wird durch die Bezugnahme auf eine deflationäre Wahrheitstheorie gewährleistet. Nach der deflationären Wahrheitstheorie ist die Hauptprämisse von Geachs Beispielargument genau dann wahr, wenn es schlecht ist, deinen kleinen Bruder dazu zu bringen, eine

Handlung auszuführen, wenn diese Handlung schlecht ist, und die Nebenprämissen von Geachs Beispielargument genau dann wahr, wenn es schlecht ist, die Katze zu quälen. Nun scheint es kein Problem für den Emotivisten zu sein, zu zeigen, dass aus diesen beiden Prämissen folgt, dass es schlecht ist, deinen kleinen Bruder dazu zu bringen, die Katze zu quälen. Schließlich folgt diese Konklusion allein aufgrund der Eigenschaften der Junktoren (zum Beispiel „wenn ..., dann ...“).

Etwas, das an diesem Lösungsansatz seltsam erscheinen mag, lässt sich mit Dreiers (1996) erfundenem Prädikat „ist hiyo“ veranschaulichen. Nehmen wir an, der Satz „Bob ist hiyo“ habe, genau wie dem Emotivismus zu Folge der Satz „Es ist schlecht, die Katze zu quälen“, eine bestimmte pragmatische Bedeutung: „Bob ist hiyo“ bedeutet pragmatisch, Bob zu begrüßen; wenn wir Bob sehen und ihn begrüßen möchten, rufen wir „Bob ist hiyo“. Betrachten wir nun das Konditional „Wenn ein Dingo in der Nähe ist, dann ist Bob hiyo.“ Dreier stellt fest, dass eine Bezugnahme auf eine deflationäre Wahrheitstheorie, der zufolge „Bob ist hiyo“ genau dann wahr ist, wenn Bob hiyo ist, nichts dazu beizutragen scheint, zu erklären, was das Konditional „Wenn ein Dingo in der Nähe ist, dann ist Bob hiyo“ bedeutet. Wir verstehen, was „Bob ist hiyo“ pragmatisch bedeutet, doch die Annahme einer deflationären Wahrheitstheorie führt nicht dazu, dass wir verstehen, was „Wenn ein Dingo in der Nähe ist, dann ist Bob hiyo“ bedeutet. In entsprechender Weise scheint Stoljars Lösungsvorschlag für den Emotivismus nicht dazu zu führen, dass wir verstehen, was „Wenn es schlecht ist, eine Handlung auszuführen, dann ist es schlecht, deinen kleinen Bruder dazu zu bringen, diese Handlung auszuführen“ bedeutet. Die Bezugnahme auf eine deflationäre Wahrheitstheorie scheint also für den Nichtkognitivismus nicht auszureichen, um die Bedeutung von eingebetteten moralischen Sätzen zu erklären. Dieses scheinbare Unvermögen stellt ein Problem für den Nichtkognitivismus dar, selbst wenn man Stoljars deflationistischen Ansatz bezüglich der Gültigkeit des von Geach angeführten Arguments akzeptiert. Das Frege-Geach-Problem bleibt somit weiterhin eine beständige Quelle von Schwierigkeiten für den Nichtkognitivismus.

Formale Detailanalyse

Es existieren verschiedene Möglichkeiten, das Argument zu formalisieren. Die nachfolgende Darstellung bietet eine vergleichsweise einfache Herangehensweise, um das Argument zu verstehen.

Es sei p: „Das folgende Argument ist gültig:

Hauptprämissen: Wenn es schlecht ist, eine Handlung auszuführen, dann ist es schlecht, deinen kleinen Bruder dazu zu bringen, diese Handlung auszuführen.
 Nebenprämissen: Es ist schlecht, die Katze zu quälen. Konklusion: Also ist es schlecht, deinen kleinen Bruder dazu zu bringen, die Katze zu quälen.“

Damit haben wir Prämisse 1 formalisiert:

1. p

Es sei a: „Die Bedeutung von „schlecht“ in der Hauptprämissen“. Es sei b: „Die Bedeutung von „schlecht“ in der Nebenprämissen“. Nun können wir Prämisse 2 wie folgt ausdrücken:

2. $p \rightarrow a=b$

In Deutsch: Wenn das obige Argument gültig ist, dann ist die Bedeutung von „schlecht“ in der Hauptprämissen identisch mit der Bedeutung von „schlecht“ in der Nebenprämissen.

Es sei c: „Das Ausdrücken einer Verurteilung“. Prämisse 3 lässt sich folgendermaßen formalisieren:

3. $a \neq c$

Auf Deutsch: Die Bedeutung von „schlecht“ in der Hauptprämissen ist nicht identisch mit dem Ausdrücken einer Verurteilung.

Es sei q: „Es ist nicht beschreibend, sondern verurteilend, einen bestimmten Akt als „schlecht“ zu bezeichnen“. Prämisse 4 lässt sich nun wie folgt formalisieren:

4. $q \rightarrow b=c$

Auf Deutsch: Wenn es nicht beschreibend, sondern verurteilend ist, einen bestimmten Akt als „schlecht“ zu bezeichnen, dann ist die Bedeutung von „schlecht“ in der Nebenprämisse identisch mit dem Ausdrücken einer Verurteilung.

Formale Ableitung

Prämissenmenge | Zeilennummer | Formel | Annotation

1 | 1 | p | Prämisse 2 | 2 | p → a=b | Prämisse 3 | 3 | a≠c | Prämisse 4 | 4 | q → b=c | Prämisse 1, 2 | 5 | a=b | 1, 2 Modus Ponens 1, 2, 3 | 6 | b≠c | 3,5 Substitution 1, 2, 3, 4 | 7 | ¬q | 4, 6 Modus Tollens

Literaturangaben

Blackburn, S. (1984). Spreading the Word - Groundings in the Philosophy of Language. Oxford University Press.

Dreier, J. (1996). Expressivist embeddings and minimalist truth. *Philosophical Studies*, 83(1), 29–51.

Frege, G. (1879). Begriffsschrift: Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Louis Neber.

Geach, P. (1957). Imperative and deontic logic. *Analysis*, 18(3), 49–56.

Geach, P. (1965). Assertion. *Philosophical Review*, 74(4), 449–465.

Geach, P. T. (1960). Ascriptivism. *Philosophical Review*, 69(2), 221–225.

Hare, R. M. (1952). The Language of Morals. Oxford University Press.

Schroeder, M. (2008). Being for: Evaluating the semantic program of expressivism. Oxford University Press.

Searle, J. R. (1962). Meaning and speech acts. *Philosophical Review*, 71(4), 423–432.

Stoljar, D. (1993). Emotivism and truth conditions. *Philosophical Studies*, 70(1), 81–101.

Unwin, N. (1999). Quasi-realism, negation and the Frege-Geach problem. *Philosophical Quarterly*, 49(196), 337–352.

Van Roojen, M. (1996). Expressivism and irrationality. *Philosophical Review*, 105(3), 311–335.