

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Національний університет «Києво-Могилянська академія»

Факультет гуманітарних наук

Кафедра філософії та релігієзнавства

Кваліфікаційна робота

(освітній ступінь – «бакалавр»)

на тему: **«Суперечність морального вибору в дилемі вагонетки»**

Виконала: студентка 4-го року навчання,
Спеціальність: 033 Філософія
Зайвая Тетяна Валентинівна

Керівник: Циба В'ячеслав Миколайович,
кандидат філософських наук, доцент

Рецензент _____

Кваліфікаційна робота захищена
з оцінкою «_____»
Секретар ЕК _____
«_____» _____ 20__ р.

Київ – 2020
ЗМІСТ

ВСТУП.....	С.3-5
РОЗДІЛ І. ІСТОРІЯ ПРОБЛЕМИ.....	С.6-13
РОЗДІЛ ІІ. ПОЗИТИВНА РОЛЬ ПРОБЛЕМИ В СИСТЕМІ МОРАЛЬНОЇ ФІЛОСОФІЇ	С.14-24
РОЗДІЛ ІІІ. НЕГАТИВНИЙ АСПЕКТ ТА КРИТИКА ПРОБЛЕМИ	С.25-31
РОЗДІЛ ІV. МОРАЛЬНИЙ ВИБІР У ШТУЧНОГО ІНТЕЛЕКТУ	С.32-38
ВИСНОВКИ.....	С.39-41
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	С.42

ВСТУП

В моїй роботі в основному фігуруватимуть невинні жертви дивних обставин. Товстуну зіштовхнуть з моста – а може й ні. На щастя, всі ці жертви є вигаданими. Однак уявні експерименти розробляються для того, аби перевірити наші інтуїтивні моральні уявлення, допомогти нам у виробленні моральних принципів, а отже, вони приносять певну практичну користь у світі, де доводиться приймати реальні рішення, і де шкода заподіюється живим людям. Сенса будь-якого уявного експерименту в етиці – виключити нерелевантні міркування, здатні заплутати наші міркування в реальних випадках. Однак, аби експеримент був корисним, він має бути подібним до реальних ситуацій.

Уявні експерименти не існують, доки їх не придумують. Книги, що розповідають про філософію, зазвичай зосереджені на ідеях, а не людях. Однак ідеї не виникають з порожнечі; вони є результатом певного часу і місця, виховання і особистих якостей. Можливо, вони були сприйняті як спростування якихось інших ідей або ж як відображення важливих для якогось конкретного моменту інтересів. Можливо, в них відображається особлива задача, яку вирішував мислитель. Є причина, з якої злочин, що є головним сюжетом цієї проблеми (а саме вбивство), так і не було в повній мірі розкрито, принаймні в філософському плані: воно складне... по-справжньому складне. Питання, які на перший погляд видаються елементарними, в міру розгляду набувають безліч додаткових відтінків і смислів.¹

Моральні судження завжди займали одне з основних місць у соціальній поведінці людини. Оцінюючи поведінку в тій чи іншій ситуації, ми вибудовуємо певні норми, які вважаємо морально правильними. Суперечність моральних виборів, які стоять перед людством кожного дня я би хотіла розглянути на прикладі найбільш поширеного морального експерименту з вагонеткою; сценарій, що у всіх випадках змушує замислитись про смерть

¹ Эдмондс Д. Убили бы вы толстяка? Задача о вагонетке: что такое хорошо и что такое плохо? / Д. Эдмондс [пер. с англ. Д.Кралечкина]. – М: Издательство Института Гайдара, 2016. – С. 10-11.

інших людей, програтися в більш ніж десяти варіаціях з різними емоційними важелями й умовами. Наразі існує настільки багато статей на рахунок морального експерименту, що виник жартівливий неологізм: «вагонеткологія» (trolleyology). Ця дилема була сформульована ще у 1967 році, але таке етичне дослідження зацікавило людей надовго – соціальні опитування і аналіз дилеми проводяться й досі; в останні роки проблема проектується на питання морального вибору у штучного інтелекту, що є досить *актуальним* аспектом питання. *Стан наукової розробки* цієї теми більш детально я проаналізую у розділі I через історію дилеми вагонетки.

Об'єктом мого дослідження є моральний вибір в дилемі вагонетки, предметом – суперечність морального вибору.

Метою мого дослідження є проаналізувати моральну дилему вагонетки та її варіації, визначити її вплив на етичний вимір, можливі наслідки і вплив дилеми на світ в майбутньому.

Моїми завданнями є:

1. Розглянути історію проблеми.
2. Проаналізувати позитивну роль проблеми в системі моральної філософії.
3. Проаналізувати негативний аспект та критику проблеми.
4. Дослідити питання моральний вибору в майбутньому – а саме машинну етику та її можливі наслідки.

Кваліфікаційна робота складається зі вступу, чотирьох розділів, висновку та списку використаних джерел. *Структура роботи* є такою: у першому розділі розкривається історія дослідження дилеми вагонетки, аналіз її різних прикладів, умов, форм. У другому розділі розкривається позитивний вплив проблеми на етичний вимір. У третьому розділі висвітлюється критика проблеми та її негативні наслідки. У четвертому розділі розкривається проекція дилеми на майбутнє, вплив її на машинну етику та подальші наслідки.

Теоретико-методологічні засади дослідження базуються на герменевтичному та компаративному методах (порівняльно-соціальний аналіз).

Джерельна база дослідження складається з комплексу статей та дослідницької літератури. Основною джерельною базою є книга Девіда Едмондса «Чи вбили б ви товстуна?», праця Філіппи Фут «Проблема абортів і доктрина подвійного ефекту», а також дослідницькі статті (Bruers S. «A review and systematization of the Trolley Problem», Сатохіна Н. І. «Проблема вагонетки і сучасні теорії справедливості», Edmonds D. «Matters of life and death» та інші статті, що зазначені в списку джерел.

РОЗДІЛ І

ІСТОРІЯ ПРОБЛЕМИ

Дилема вагонетки складається з низки випадків, пов'язаних з візком, що загрожує життю певної кількості людей. Основна структура всіх прикладів дилеми однакова: якщо не діяти, загине п'ять людей; якщо діяти, одна людина буде вбита, а п'ятеро – врятовані. Дослідження щодо цього етичного питання показали, що інтуїція більшості людей не відповідає ні чистій деонтології, ні утилітаризму. Під «чистою» деонтологією ми маємо на увазі те, що люди повинні дотримуватися наступного правила: ніколи не дійте, якщо цей акт завдасть шкоди людям, яким вона не загрожувала, якщо ви не діяли. Під «чистим» утилітаризмом я маю на увазі те, що люди повинні дотримуватися цього правила: завжди обирайте дію, яка максимально збільшує кількість врятованих життів (тобто, найменше загальної шкоди).²

Різні дилеми вагонеток мають однакову послідовну структуру, але все-таки люди, що зіткнулися з цими дилемами навряд чи скажуть, що ніколи не слід діяти або що завжди слід діяти. Коли виникають різні дилеми, більшість людей кажуть, що в одному випадку ми повинні діяти, але в іншій дилемі нам заборонено діяти; це так, ніби люди роблять невідповідний, суперечливий вибір. Лише чистий утилітарний консеквенціалізм³ стверджує, що ми завжди повинні діяти у всіх випадках. Тож моральні інтуїції більшості людей відхиляються від цієї консеквенціалістичної етики, і тому проблема вагонетки є цікавим мислительним експериментом для вивчення деонтологічної етики. Основне питання полягає в наступному: яка морально важлива різниця між Дилемою А і Дилемою Б – як морально допустимо діяти в одній ситуації вибору, але не допустимо в іншій? Ми розуміємо, що послідовне моральне розв'язання проблеми вагонетки повинно містити чіткий опис правил чи

² Bruers S. A review and systematization of the Trolley Problem / S. Bruers, J. Braeckman // Springer Science. – 2013. – Режим доступу: <https://rdcu.be/b345Z> С.252

³ Консеквенціалізм, як випливає з назви, є думкою, що нормативні властивості залежать лише від наслідків. Ця історично важлива і все ще популярна теорія втілює таку інтуїцію: правильним є те, що робить світ кращим у майбутньому, оскільки ми не можемо змінити минуле. Consequentialism. Stanford Encyclopedia of Philosophy – Режим доступу: <https://plato.stanford.edu/entries/consequentialism/>

принципів, які найкращим чином відповідали б, обґрунтовували та пояснювали різноманітність моральних інтуїцій людей у різних випадках. Іншими словами, найкращим розв'язанням проблеми вагонетки є чіткий алгоритм, який вирішує, як саме слід діяти, а відповіді, які цей алгоритм генерує, повинні відповідати інтуїції.⁴

Я коротко представлю та систематизую найбільш дискутовані версії дилеми вагонетки:

Дилема №1: Перемикач. На основній колії тролейбус рухається у напрямку п'яти осіб. Ви стоїте біля перемикача. Якщо ви повернете перемикач, тролейбус буде переведений на бічну колію, але на цій бічній колії знаходиться одна людина. Якщо повернути перемикач, це призведе до смерті цієї людини, але п'ять людей, що перебувають на головному треку, будуть врятовані. Чи слід повернути вимикач? Більшість людей (приблизно 90% за даними Хаузера у 2008⁵) кажуть, що вам це дозволяється.

Дилема 2: Міст. Людина з зайвою вагою (товстун?) стоїть на мосту над доріжкою. Ви можете врятувати п'ятьох людей на доріжці внизу, штовхнувши цю людину з мосту перед вагонеткою, щоб вагонетка була зупинена його великою вагою. Людина загине, а п'ятеро людей врятуються. Лише кілька людей (приблизно 10% за даними Хаузера) стверджують, що нам дозволяється штовхнути людину з зайвою вагою. Решта респондентів або відмовляються підштовхувати, або засуджують цю дію. За словами Вальдмана та Дітериха⁶, люди більш толерантні до того, аби виштовхувати когось на колії, але у своєму дослідженні дилеми не передбачали близькості та особистого контакту з жертвою, яку підштовхують.

⁴ Bruers S. A review and systematization of the Trolley Problem / S. Bruers, J. Braeckman // Springer Science. – 2013. – Режим доступу: <https://rdcu.be/b345Z> C.253

⁵ Hauser M. Reviving Rawls' linguistic analogy: operative principles and the causal structure of moral actions [Електронний ресурс] / M. Hauser, L. Young, F. Cushman // Moral Psychology and Biology, Oxford U. Press, NY. – 2008. – Режим доступу: https://www.academia.edu/2050790/Reviving_Rawls_linguistic_analogy

⁶ Waldmann M. Throwing a bomb on a person versus throwing a person on a bomb. Intervention myopia in moral intuitions [Електронний ресурс] / M. Waldmann, J. Dieterich // Psychological Science. – 2007. – Режим доступу: https://www.psych.uni-goettingen.de/de/cognition/publikationen-dateien-waldmann/2007_bomb.pdf

Дилема 3: Петля. Як і в першій дилемі, ви стоїте біля перемикача. Але цього разу бічна доріжка з'єднана з основною колією. Якщо на бічній колії немає нікого, вагонетка все одно рухатиметься на основну колію і вб'є п'ятьох людей. Але на бічній колії – людина з зайвою вагою. Тож якщо ви повернете перемикач, людина з зайвою вагою перекриє шлях вагонетці. Загине вона, врятується п'ятеро людей. В недавньому опитуванні Хаузера⁷ виявилось, що ледве половина респондентів дозволили повернути перемикач. Однак, на думку Вальдмана та Дітериха⁸, люди більш толерантні до повороту перемикача. Вони побудували дилему по-іншому, де людина з зайвою вагою – на бічній колії сидить в автобусі. Тож на бічній колії автобус перекриє вагонетку, а не людина. Людина в автобусі загине в аварії.

Дилема 4: Петля з каменем. Ситуація нагадує ситуацію в Дилемі 3. Але цього разу важкий камінь розташований позаду людини на бічній колії. Чоловік на бічній колії недостатньо важкий, щоб зупинити вагонетку. Камінь перекриє вагонетку, якщо повернути перемикач. Але людина, що знаходиться перед каменем загине. В опитуванні, проведеному Хаузером⁹ було встановлено, що майже три чверті респондентів кажуть, що дозволено повернути перемикач, і це є статистично значущою відмінністю від Дилеми 3.¹⁰

Дилема 5: Вантажівка. Можна заблокувати вагонетку, штовхнувши важку вантажівку на рейки. У цій вантажівці є один пасажир, який загине (цю дилему вивчали Вальдманн і Дітерих¹¹).

⁷ Hauser M. Reviving Rawls' linguistic analogy: operative principles and the causal structure of moral actions [Електронний ресурс] / M. Hauser, L. Young, F. Cushman // Moral Psychology and Biology, Oxford U. Press, NY. – 2008. – Режим доступу: https://www.academia.edu/2050790/Reviving_Rawls_linguistic_analogy

⁸ Waldmann M. Throwing a bomb on a person versus throwing a person on a bomb. Intervention myopia in moral intuitions [Електронний ресурс] / M. Waldmann, J. Dieterich // Psychological Science. – 2007. – Режим доступу: https://www.psych.uni-goettingen.de/de/cognition/publikationen-dateien-waldmann/2007_bomb.pdf

⁹ Hauser M. Reviving Rawls' linguistic analogy: operative principles and the causal structure of moral actions [Електронний ресурс] / M. Hauser, L. Young, F. Cushman // Moral Psychology and Biology, Oxford U. Press, NY. – 2008. – Режим доступу: https://www.academia.edu/2050790/Reviving_Rawls_linguistic_analogy

¹⁰ Bruers S. A review and systematization of the Trolley Problem / S. Bruers, J. Braeckman // Springer Science. – 2013. – Режим доступу: <https://rdcu.be/b345Z> C.254

¹¹ Waldmann M. Throwing a bomb on a person versus throwing a person on a bomb. Intervention myopia in moral intuitions [Електронний ресурс] / M. Waldmann, J. Dieterich // Psychological Science. – 2007. – Режим доступу: https://www.psych.uni-goettingen.de/de/cognition/publikationen-dateien-waldmann/2007_bomb.pdf

Дилема 6: Камінь. Ви можете повернути перемикач, перенаправляючи вагонетку на бічну колію. На цій бічній колії є великий камінь. Коли вагонетка потрапляє на камінь, камінь сковзає у сторону і вбиває людину.

Дилема 7: Платформа. П'ять людей перебувають на рухомій платформі на рейках. Якщо нічого не робити, вагонетка розгромить платформу і вб'є п'ятьох. Але ви можете перемістити платформу від рейок, аби врятувати людей. Але цей хід підштовхне іншу людину (яка стоїть поруч із платформою) на електричний кабель. Отже, ця людина помре від електричного струму.

Дилема 8: Петля з лавиною. Ця дилема схожа на дилему петлі, але людина на бічній колії контролює бар'єр безпеки перед лавиною. Людина на бічній колії недостатньо важка, аби зупинити вагонетку, тому потрібно створити лавину. Але лише після того, як людина помре, можна ініціювати лавину, яка не перекрита бар'єром; тому лише після того, як жертва помирає і більше не контролює бар'єр, візок може бути заблокований, а п'ятьох врятовано.¹²

Філіппою Фут в праці «Проблема абортів та доктрина подвійного ефекту»¹³ задача була сформульована так: «Філософи добре знають історію про товстуна, що застряг у виході з печери. Декілька мандрівників йшли за товстуном, який застряг, ніби-то захоплюючи їх і не випускаючи з печери. Звісно, правильним було б почекати, поки огрядна людина схудне. Але філософи побачили, що у печері почала підійматися вода. На щастя (на щастя?) в печері знайшовся динаміт, яким можна підірвати людину та відкрити вихід з печери. В цій ситуації мандрівникам залишається або використати динаміт, або потонути в печері. В одній версії товстун потоне разом з усіма, в іншій – всі виживуть та виберуться разом. Але проблема полягає ось в чому – чи слід використовувати динаміт?»¹⁴. Ситуація непроста,

¹² Bruers S. A review and systematization of the Trolley Problem / S. Bruers, J. Braeckman // Springer Science. – 2013. – Режим доступу: <https://rdcu.be/b345Z> С.254

¹³ Foot P. The Problem of Abortion and the Doctrine of the Double Effect / Philippa Foot // Oxford Review, No. 5. [Електронний ресурс] – 1967. – Режим доступу: <https://philpapers.org/archive/footpo-2.pdf> Р. 1-2

¹⁴ Цит. за Foot P. The Problem of Abortion and the Doctrine of the Double Effect / Philippa Foot // Oxford Review, No. 5. [Електронний ресурс] – 1967. – Режим доступу: <https://philpapers.org/archive/footpo-2.pdf> Р. 1-2

тому що моральний вибір індивіда пов'язаний з вбивством іншої людини. Тобто, необхідно діяти активно та навмисно, хоча і заради благої мети.

Джозеф Томпсон також запропонував цікаву альтернативу цій задачі: хірург-трансплантолог має п'ять пацієнтів, що потребують трансплантації різних органів, без пересадки яких на кожного чекає загибель. На превеликий жаль, для здійснення цих пересадок немає доступних органів. Молодий турист, що зупинився у цьому місті, прийшов на звичайний медогляд. За час медогляду, хірург розуміє, що всі органи туриста підходять для усіх п'яти пацієнтів, що помирають. Якщо уявити, що турист зникне, хірурга ніхто підозрювати не буде. Чи можливо в такому випадку хірургові вбити туриста (особливістю є безкарність дії) заради порятунку п'яти хворих?¹⁵

Проблема вагонетки та сучасна теорія справедливості постають у цікавому взаємовідношенні, що має логічно обґрунтовуватись. Розглянемо вибір індивіда у наступних аспектах: кантіанства та утилітаризму. Іммануїл Кант був засновником деонтологічної етики, що своєю чергою завбачає: коли ми визнаємо невід'ємними права людини у будь-якому випадку (право на життя включно) – ми не можемо позбавляти життя навіть одну людину заради порятунку декількох. Якщо йти саме за таким напрямком думки, це є етикою мотиву, в тому сенсі, що зміст має не вчинок чи його наслідки, а безпосередньо підстави воління, на яких вчиняється певна дія. У результаті, той факт, що моральний закон був реалізований є більш важливим фактором, ніж наслідки, до яких це призвело.¹⁶

Нарешті, з позиції етики обов'язку, моральна теорія, що допускає реальний конфлікт обов'язків, має розглядатися як недостатньо опрацьована або помилкова. Ця позиція спирається на категоричний імператив І. Канта – на абсолютному, моральному міркуванні, що людина не може стати способом розв'язання проблеми, а отже, ми не маємо права використовувати одного

¹⁵ Сатохіна Н. І. «Проблема вагонетки» і сучасні теорії справедливості / Н. І. Сатохіна, Л. А. Улашкевич. // Молодий вчений. – 2018. – С.2

¹⁶ Там само. С.2

індивіда навіть заради порятунку п'ятьох людей.¹⁷ Втім варто зазначити, що можливим є варіант самопожертви заради порятунку всіх. Тобто вибір можна зробити на користь самопожертви, але не за іншу людину (і пожертвувати нею). Будь-яке життя, не дивлячись на те, чи це є злочинець, чи винахідник вакцини від усіх хвороб, є цінним. Через це неможливо ставити його за предмет торгу в моральній дилемі.¹⁸

Утилітаристська концепція пропонує нам наступний варіант: потрібно відкинути питання моралі та обрати шлях спричинення найменшого зла. Основним аспектом тут виступає кількість (жертв). Таку концепцію захищати простіше, оскільки ми ніби виправдовуємо жертву однієї людини добрими намірами. Пояснення тут доволі доступне і просте – завжди варто запобігати злу. Смерть однієї людини є меншим злом ніж зло, що потягне за собою п'ять смертей. Звичайно, в окремих ситуаціях людина намагатиметься вчинити якнайкраще, з добрих намірів. Безперечно, аби проігнорувати такий порив, потрібна певна сміливість та воління взяти відповідальність за власний вибір заради універсальних, загальнолюдських моральних принципів, дотримуватися яких нам безумовно приписує наш розум (деонтологічна етика); тобто людина повинна взорувати свої вчинки не на користь власних переконань, а узгоджувати їх із тими засадами, які зорієнтовують її у проблемних випадках, які є моментом «зустрічі» загального й конкретного, універсального, етичного та персонального, прагматичного. Але ж враховуючи доктрину подвійного ефекту, ми можемо морально виправдати дію, що веде до негативних наслідків, якщо дотримуватись певних умов:

1. Задум отримати позитивний результат – отримане добро може слугувати виправданню зла, яке було заподіяне.

¹⁷ За категоричним імперативом Канта стоїть певне роз'яснення, чому людина не може бути засобом досягнення моральної мети: людина має чинити згідно з вимогами морального закону та відноситися до інших людей не шукаючи вигоди, завбачаючи не засіб досягнення певної цілі, а тільки абсолютну самостійну самоцінність особистості (тобто людина є вищою цінністю).

¹⁸ Сатохіна Н. І. «Проблема вагонетки» і сучасні теорії справедливості / Н. І. Сатохіна, Л. А. Улашкевич. // Молодий вчений. – 2018. – С.2

2. Дія має бути морально нейтральною (чи позитивною), але ні в якому разі не негативною. Безвідносно до результату, дія не може бути негативною.
3. Позитивний результат має бути прямим наслідком самої дії, а не негативних наслідків. Добро не може бути наслідком того, що спочатку скоїло зло.
4. Негативне не має переважати над позитивним. Якщо основними є позитивні наміри, але отримана шкода є більшою, умова вважається недотриманою.¹⁹

Різноманітні варіанти дилеми вагонетки та особливості кожної людини, що стоїть перед вибором, і спонукають до дії або бездіяльності на користь утилітаризму чи деонтології. Дехто з нас може виправдати свою провину, а дехто не може взяти на себе відповідальність за певний вчинок. У такий спосіб, можна подумати, що кантіанство може бути зумовлене боягузтвом, а утилітаризм – оправдовуванням на свою користь. Нарешті, кожна з цих версій має свої аргументи на користь доведення власної справедливості. Але не можна сказати, що якась з них є справді істиною у своєму підсумку. Як і всі концепції, вони мають переваги й недоліки. Кантіанство ставить перед нами сценарій пожертви п'ятьма індивідами, дотримуючись цінності чужого людського життя і права вирішувати долю інших людей. Утилітаризм, своєю чергою, пропонує піти шляхом найменшого зла, що також є раціональним. Саме тому кожна з цих концепцій може відображатися у поглядах людей, які будуть обирати, який з цих варіантів є близьким та правильним. Звичайно, важно переоцінити вплив дилеми вагонетки в умовах постійного розвитку етики та моралі. З часом людина від чогось відмовляється, щось знаходить, а до чогось повертається. Так от, саме моральні дилеми і є основою суспільної свідомості, яку потрібно постійно переосмислювати та намагатися дійти згоди при її вирішенні. За останніх років, дилема вагонетки знайшла своє місце не

¹⁹ Сатохіна Н. І. «Проблема вагонетки» і сучасні теорії справедливості / Н. І. Сатохіна, Л. А. Улашкевич. // Молодий вчений. – 2018. – С.3

тільки в рамках філософії, а й у сфері психології, когнітивної науки та нейроетики²⁰. Зокрема, в експериментальній філософії проблема вивчає інтуїцію звичайних людей за допомогою проведення різних опитувань. Прояснення такої інтуїції дозволяє зрозуміти, наскільки теорії та концепції філософів близькі до уявлень звичайних людей.²¹

Для нейроетики моральні дилеми відіграють особливу роль: багато моральних (і не тільки) поведінкових рішень, зумовлені несвідомими аспектами, які зображають певні нейрофізіологічні особливості функціонування нашого мозку. Наприклад, наші моральні рішення тісно пов'язані зі здатністю людини до емпатії; з нейрофізіологічної ж сторони, емпатія зумовлена активністю окремих ділянок мозку. Скажімо, в ситуації з вагонеткою, людина більш схильна до емпатії поставить себе на місце жертв та на основі своїх відчуттів зробить вибір. Когнітивна наука, об'єднує в собі декілька наук, таких як психологія, теорія пізнання, нейрофізіологія тощо. Для прикладу, як змінюється вибір людини, якщо вона розуміє, що доведеться пожертвувати рідною або близькою людиною? Чи зміниться вибір, якщо виживання інших не гарантоване? Відповідь може бути і позитивною, і негативною. Дехто може ризикнути, а хтось ставить менше зло вище за близькі стосунки. Відповідь на такі питання завжди матиме індивідуальний характер.²²

Отже, у цьому розділі було розглянуто історію дилеми вагонетки та її різноманітні варіанти, що утворились впродовж дослідження проблеми. Було згадано як і про початковий варіант, вперше описаний Філіппою Фут, так і про інші більш складні ситуації, що супроводжувались соціологічними дослідженнями. Також був проаналізований аспект суперечності деонтології (а саме кантіанської етики) та утилітаризму в моральній дилемі.

²⁰ Про нейроетику та її вплив на дилему вагонетки більш детально розкрито у Розділі III.

²¹ Сатохіна Н. І. «Проблема вагонетки» і сучасні теорії справедливості / Н. І. Сатохіна, Л. А. Улашкевич. // Молодий вчений. – 2018. – С.3

²² Там само. С.3

РОЗДІЛ II

ПОЗИТИВНА РОЛЬ ПРОБЛЕМИ В СИСТЕМІ МОРАЛЬНОЇ ФІЛОСОФІЇ

Історія цього морального експерименту і справді є цікавою не тільки для філософів чи психологів – через свій азарт і різні сценарії суперечності щодо морального вибору можуть виникнути у будь-якої людини. Ця дилема стала настільки актуальною та значущою, що захопила навіть телевізійний та інтернет-простір.²³ Та чи є ця проблема настільки корисною для нашого мозку, наскільки вона є цікавою? І чим може допомогти розуміння цієї проблеми у реальному житті?

Як демонструють залучені нами дані, люди скоріше принесуть в жертву одну людину, але не товстуна. Однак в обох з цих випадків одна людина помирає, аби врятувати п'ять інших. У чому ж полягає відповідне етичне розмежування між ними? Це питання породило процвітаючу академічну міні-індустрію, що називається тролейологія. Тролейологія охоплює найглибшу напруженість у нашому моральному світогляді. Аби захоплювати нашу моральну інтуїцію, філософи вигадували все більш геніальні сценарії. Тролейбус, як правило, біжить до п'яти нещасних, і читачеві пропонують різні засоби, аби врятувати їх ціною іншого життя, залучаючи такі реквізити як товсті джентльмени, пішохідні мости та пастки. Але в корені проблеми вагонетки – філософська інтрига, яка приваблює найрозумніші уми моральної філософії.²⁴

Один з них – Джефф МакМахан з Ратгерського університету. Мак-Махан привертає увагу до витончених формулювань в моральній філософії. Він вважає, що проблема вагонетки надає важливості вченню, вперше створеному в XIII столітті Томою Аквінським. Аквінат склав принципи, необхідні для

²³ Наприклад, висвітлення дилеми вагонетки відбувається у серіалі The Good Place (Netflix, Season 2, Ep.5), а власники ютуб-каналу Mind Field вирішили провести експеримент з реальними людьми (Season 2, Ep. 1). На щастя, вибір між тим кого вбити був не справжній – у епізоді досліджується реакція людини, а після експерименту з кожним учасником проводилась пост-травматична психологічна сесія. Mind Field (Season 2, Ep. 1) – Режим доступу: <https://youtu.be/1sl5KJ69qiA>

²⁴ Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

справедливої війни, і був першим мислителем, який виклав вчення про подвійний ефект, наріжний камінь не просто католицької етики, але й моралі здорового глузду. Грубо кажучи, вчення легітимує здійснення вчинку, що має певні погані наслідки, якщо напротивагу вчинок є хорошим і якщо погані наслідки – ненавмисними. Застосовуючи це вчення до проблеми з вагонеткою, стверджувалося, що в першому сценарії його суб'єкти не мають наміру вбити одну людину. Якби ви повернули перемикач та запустили вагонетку, а людина дивом врятувалася – ви були б у захваті. Але у другому сценарії ви маєте намір вбити товстуна. Якби він відскочив від колії та втік зі шляху вагонетки, це зірве вашу мету, бо п'ять людей все одно будуть вбиті.²⁵

Макмахан іноді відвідує федеральну академію Вест-Пойнт штату Нью-Йорк, аби прочитати лекцію про дилему вагонетки та інші аспекти прикладної етики. Студенти тут не середньостатистичні студенти коледжу – в академії готують курсантів офіцерів США. Ці юнаки та дівчата є майбутніми лідерами найпотужнішої військової сили у світі. Всі курсанти Вест-Пойнт проходять обов'язковий курс з філософії та теорії «справедливої війни», що включає дилему вагонетки. Девід Едмондз розмовляв з кількома винятковими курсантами, які були ознайомлені з дилемою вагонетки, і ніхто з них не заперечив би вбивці товстуна. Вони пояснили, що ці два сценарії зображають різницю між націленням на військову установу, знаючи, що цивільне населення буде вбито, і свідомим вбивством цивільних.²⁶

Цих курсантів вчать приймати спонтанні рішення (що містять моральну оцінку) в екстремальних ситуаціях бойових дій самостійно, а не сліпо слідувати правилам. Один з інструкторів, майор Денні Казьєр, визнав це, але сказав, що «для успішного завершення академії це зависока ціна» (the pay-off is too high to pass on). Він каже, що надважливо, коли солдати опиняються в жахливій ситуації на полі бою, аби вони не втрачали з поля зору «основні

²⁵ Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

²⁶ Там само.

принципи, в які людина вірить, і які керують її діями. І ці принципи повинні бути обумовлені такими міркуваннями, як дилема вагонетки». Курсанти погоджуються.²⁷

Вчення про подвійний ефект закріплене в Женевських конвенціях.²⁸ Ця дилема також має вирішальне значення для інших сфер моралі, таких як медична етика. У західних країнах існує чітке розмежування між швидким вбивством хворого пацієнта та надаванням ліків для зняття болю, усвідомлюючи, що це прискорить його смерть. Майкл Роулінс, голова Національного інституту здоров'я і досконалості допомоги, наводить старе прислів'я та виділяє відмінність, пов'язану з доктриною подвійного ефекту, що знаходиться між діянням і упушенням: «Не вбивай, але й не рятуй життя попри волю його власника»²⁹ (Thou shall not kill but need not strive officiously to keep alive).

Коаліція стверджує, що фінансування Національної служби охорони здоров'я буде обмежено, а ресурси мають бути скорочені в найближчі кілька років. Національний інститут здоров'я і досконалості допомоги, який розробляє вказівки щодо лікування для Національної служби охорони здоров'я, зіткнеться з інтенсивним оглядом. Так само буде розроблений алгоритм, який визначає, чи варто лікування його ціни, враховуючи вартість збережених життів та їх якість. «Коли одна пацієнтка погодилась на [препарат] герцептин, для раннього раку молочної залози, один відділ Національної служби охорони здоров'я закрит свою діабетичну клініку, аби заплатити за це», – сказав Роулінс. «Це жахливі рішення, які потрібно прийняти».³⁰

Але, всупереч впливу, який вчення про подвійний ефект має на політику, воно може не пояснити всі наші інтуїтивні відповіді на дилему. Розгляньмо

²⁷ Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

²⁸ Protocol additional to the Geneva Conventions of 12 August 1949, and relating to the protection of victims of international armed conflicts (Protocol I) (adopted 8 June 1977, entered into force 7 December 1978). Режим доступу: <https://treaties.un.org/doc/Publication/UNTS/Volume%201125/volume-1125-I-17512-English.pdf>

²⁹ Цит. за Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

³⁰ Там само.

сценарій Дилеми №3. Вагонетка збирається вбити п'ятьох людей, але її можна повернути по іншій колії, до якої прив'язаний велика людина. Зіткнення з однією людиною зупинить візок. Чи слід повернути перемикач? У дослідженнях більшість людей вважають, що було б правильно повернути. Але якби людина вирвалася на волю та втекла з колії, п'ятеро людей все одно загинули. Так само, як і в Дилемі №2, ви потребуєте смерті однієї людини, аби врятувати п'ятьох.³¹

Прибічники утилітаризму, що вірять у велике благо, підозріло ставляться до цих експериментів. Нам потрібно подолати наші ірраціональні бажання і скинути товстуна, кажуть вони. Але Камм³² сприймає серйозно окремість людей. Вона вважає, що ми не можемо спричинити людям деякі речі (наприклад, катування), попри те, чи сприятиме це загальній сумі щастя. І вона захищає себе від критиків, які звинувачують її у створенні ризикованих сценаріїв, які не мають стосунку до реальності. «У реальних ситуаціях відбувається багато факторів, і важко перевірити, чи є цей фактор найважливішим, чи інший. Ви повинні штучно конструювати випадки, аби зосередити увагу на важливих факторах. Як учений в лабораторії, який повинен з'ясувати, чи, скажімо, частинка пилу впливає на тертя, і намагається утримувати все інше постійним».³³

Дилема вагонетки не є академічною вправою для Камм. Їй помітно неприємна думка про те, що люди інструменталізуються, їх інтереси неупереджено зважуються в якомусь клінічному утилітарному алгоритмі. І все ж вона любить своєрідний світ, населений гіпотетичними та сюрреалістичними моральними дилемами. «Я відчуваю, що мене прийняли до

³¹ Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

³² Френсіс Камм (Frances M. Kamm) – американська філософиня, що спеціалізується на нормативній та прикладній етиці. Професорка філософії та державної політики в Гарвардському університеті. Режим доступу: <https://www.hks.harvard.edu/faculty/frances-kamm>

³³ Цит. за Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

цілого світу відмінностей, яких не бачили інші. І я захоплена цим світом, неначе знаходжуся у прекрасній картині».³⁴

Більшість людей вважають, що правильно повернути важіль. Більшість людей вважають, що невірно підштовхнути товстуна. Що дивно – ніхто із нас насправді не стикався з цим становищем.

Оскільки філософи досі не здатні досягти консенсусу з дилемою, їхні більш емпірично налаштовані колеги з інших дисциплін замислилися. Наразі було проведено багато досліджень щодо нашої інтуїції та прийняття рішень, і вони, схоже, не сильно різняться між чоловіками та жінками, багатими та бідними, старими та молодими, між релігійними та світськими, між докторами та тими, хто не має кваліфікації.³⁵

Гарвардський психолог Марк Хаузер вже кілька років проводить інтернет-опитування під назвою Moral Sense Test³⁶. Будь-хто може скласти цей тест, що доступний декількома мовами. Це, своєю чергою, створює низку моральних труднощів. Ноам Чомський стверджував³⁷, що всі мови мають спільну глибоку структуру і що мовний інстинкт вроджений. Хаузер висуває подібну тезу щодо моральності. Він каже, що у нас є вроджене відчуття вкрай абстрактних правил. У них немає конкретного вмісту, наприклад, «не вбивай матусю», але, однак, існують абстрактні правила щодо того, коли вбивство чи заподіяння шкоди допустимо. Хаузер вважає, що вчення про подвійний ефект, вперше викладене майже тисячоліття тому, є провідним для нас.

Носій англійської мови радше за все скаже «Великий, білий потяг», а не «білий, великий потяг», але, ймовірно, не зможе дати вам миттєвого пояснення, чому. В мові, глибинна структура є певним засобом опису структури речення, що утворено загальними елементами смислу. Засновником теорії глибинних структур якраз і був Н. Чомський: він запевняв, що глибинна

³⁴ Цит. за Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

³⁵ Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

³⁶ Moral Sense Test. Режим доступу: <http://www.moralsensetest.com/>

³⁷ Chomsky N. The 'Chomskyan Era' / Noam Chomsky // The Architecture of Language. – 2000. – Режим доступу до ресурсу: <https://chomsky.info/architecture01/>

структура є загальним змістом речення.³⁸ Так само, стверджує Хаузер, ми дотримуємось граматики моралі, яку ми часто не можемо сформулювати, але яка може бути з багатьма нюансами та шарами. Припустімо, ми скорегували дилему вагонетки. Уявіть, що п'ять людей на трасі страждали від жахливої хвороби та все одно скоро помруть. Або уявіть, що ми виявили, що одного чоловіка прив'язали до колії п'ять хуліганів, які згодом потрапили в пастку на головній лінії. У цих випадках наші інтуїції (передзнання) щодо того, чи слід повертати перемикач, навряд чи будуть настільки безпечними. Хаузер зазначає, що його величезна база даних про глобальні моральні інстинкти дозволяє йому розробити широку картину нашої складної моральної граматики, «яка придатна для наших миттєвих, часто автоматичних рішень про те, що є правильно і що неправильно».³⁹

Хаузер – суперечлива постать, яка щойно була визнана винною у вчиненні ним наукових проступків (звинувачення висунув університет, у якому він працював). Зараз він у невизначеній відпустці. Двома поверхами вище кабінету Хаузера знаходиться кабінет іншого філософа – Джоша Гріна, молодого, кучерявого професора, який навчався на психології та філософії. Як і інші науковці, чия робота пов'язана з боротьбою з травматичними (навіть якщо й гіпотетичними) дилемами, він невгамовно веселий. Його студентка переймається щодо наступного питання. Уявіть собі: саме під час другої світової війни ви ховаєтесь з іншими від нацистів. У вас на руках голодне немовля. Якщо затиснути йому рота – воно задихнеться, але виживуть всі ті, хто переховуються разом із вами; не затиснути – дитина виживе, але всіх буде виявлено – і врешті дитя загине разом з вами. Як варто учинити?

«Я не могла б цього зробити», – каже студентка. «Я не могла б вбити свою дитину». Але шум призведе до смерті всіх. «Це не має значення. Я не вб'ю свою дитину. Я вважаю за краще, щоб це було на їх совісті». Але ж дитина

³⁸ Архелюк В. В. Актуальні проблеми структурної лінгвістики / В. В. Архелюк. // Молодий вчений. – 2015. – С. 177

³⁹ Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

все одно буде мертвою через кілька хвилин? «Так, але не від руки матері» – відповідає студентка.

Грін помістив індивідів у МРТ-сканери, коли вони розмірковували про моральні дилеми. Він вважає, що те, що відбувається, є свого роду поєднанням нейронної боротьби між обчислювальною частиною мозку, яка максимізує витрати та вигоди (краще вбити одну людину та врятувати п'ять інших) та емоційною частиною (не можна вбити власну дитину). Або, як він вважає за краще, між ефективністю та гнучкістю частин мозку. Люди з пошкодженням вентромедіальної префронтальної кори (тієї частини головного мозку, в якій емоції подаються для прийняття рішень), набагато більш схильні до вбивства товстуна. В одному дослідженні пацієнтів з цим пошкодженням виявилось приблизно в п'ять разів більше, ніж звичайних людей; всі вони погодились, що допустимо штовхнути товстуна.⁴⁰

Для тих, хто не має пошкоджень мозку, Грін провів аналогію з камерою. Камера має автоматичні налаштування, що є корисним, оскільки економить час. Але часто хочеться зробити щось цікаве і особливе, наприклад, з фокусом збоку і з розмитим фоном. Тут потрібен ручний режим. «Емоційні відповіді – це як автоматичні налаштування у вашій камері. Гнучка форма планування дій, це ручний режим». Еволюція наділила нас гнучкою системою для проблем, які ми не можемо передбачити.

Його претензії суперечливі, але він говорить, що в ситуації з «товстою людиною» є два фактори, які стоять за відразою вбити. «Вони не мають жодного або мають невеликий вплив окремо, але якщо ви комбінуйте їх, вони створюють ефект, який набагато більший, ніж сума окремих ефектів. Це схоже на взаємодію з наркотиками: якщо ти вживаєш наркотик А – тобі добре, якщо вживаєш наркотик В – тобі добре, але якщо ти приймеш їх разом – буде щось приголомшливе!»⁴¹

⁴⁰ Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

⁴¹ Цит. за Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

Одним із двох аспектів є вчення про подвійний ефект; ми піклуємось скоріше про наміри, ніж про непередбачені ефекти. Інша справа – особиста сила. Це різниця між безпосереднім фізичним впливом – наприклад, штовханням – і натисканням на перемикач. Дослідження Гріна дають змогу припустити, що люди готові вбити товстуна, якщо вони зможуть це зробити повернувши важіль чи підставивши пастку. Це чудово висвітлює відмінний підхід філософа та психолога – психологічне небажання вбивати може змінюватись залежно від того, чи це підштовхування, чи натискання на перемикач – але філософ хоче відкинути будь-яке моральне розмежування між цим.⁴²

Грін стверджує, що щодо наших відповідей на дилему вагонетки існують прямолінійні еволюційні пояснення. «Ми еволюціювали в середовищі, в якому ми досить напругу використовували наше тіло. Зараз ми живемо у світі технологічно опосередкованих дій. Товстун викликає своєрідну емоційну тривогу, яка говорить: «Ей, не робіть цього» ». Якщо Грін має рацію щодо наших інстинктів на рахунок насильства, це має великі наслідки для майбутнього ведення війни, яка все більше характеризуватиметься натисканням перемикачів, а не штовханням.⁴³

Можна було б давно передбачити, що сценарії вагонетки рано чи пізно закінчуватимуться, але виявляється, що для охочих стрибнути у вагонетку є численні галузі розслідування. Одне дослідження ставило запитання суб'єктів, чи не штовхнуть товсту мавпу з моста, аби врятувати п'ять інших мавп (відповідь «так»: ми менш деонтологічні, коли маємо справу з іншими видами). Інше дослідження назвало товстуна, перевіряючи, чи може наша реакція піддатись впливу через расу. Суб'єктам був запропонований вибір між штовханням «Тірона Пейтона» (поширене прозвисько афроамериканців) з пішохідного мосту, аби врятувати 100 членів Нью-Йоркської філармонії та

⁴² Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

⁴³ Там само.

штовханням «Чіп Еллсворт III», щоб зберегти 100 членів джазового оркестру «Гарлем». Дослідники виявили, що консерватори були байдужими між цих варіантів, але, з боку лібералів бідний старий Чіп справлявся не так добре як Тайрон.⁴⁴

У багатьох філософів метастази тролейології викликають гнів. Тролейофіли вірять у цінність таких мислительних експериментів та задовольняють винахідливістю своїх прикладів та місць, де їх можна застосувати. Але противники почуваються набагато сильніше – вони зневажають те, що представляють, і хотіли б перекинути вагонетки у віддалене пенсійне депо. Один чудовий філософ сказав: «Вибачте, я просто не працюю з вагонетками».⁴⁵

Найзапекліші противники вважають, що весь цей підхід до етики є глибоко невірним і, найбільш фундаментальним чином, неправильно характеризує природу моралі. Світ занадто складний, судження занадто багатогранні, а властивості чесноти та мудрості занадто неоднозначні, щоб ми відділяли інтуїцію зі сценаріїв вагонеток і корисно переносили їх у реальний світ. Думка полягає в тому, що важко знати, як застосовувати моральну філософію будь-яким іншим способом. Якщо і справді допустимо вбити одну людину, але неправильно вбити товстуна, нам потрібно розплутати принципи, про які йде мова. Джудіт Джарвіс Томсон якось назвала проблему вагонетки як «прекрасна, неприємна складність». Розв'язання цієї неприємної проблеми є корисним для того як ми будемо розглядати дії, що є вирішальними для життя. Це, буквально, питання життя та смерті.⁴⁶

Ось приклад: 13 червня 1944 року нацисти оприлюднили нову зброю – летючу бомбу. Звук лякав лондонців: вона гула як збентежена бджола, а потім все моторошно мовчало. Тиша сигналізувала про те, що у неї закінчилося пальне і вона летить на землю. При контакті з землею вона спричиняла

⁴⁴ Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

⁴⁵ Там само.

⁴⁶ Там само.

руйнівний вибух, миттєво руйнувавши будівлі до уламків. Нацисти називали їх ракетами Фау-1. Вони планували націлити ракети на центр столиці, яка була густо заселена і містила інститути влади. Але, німецьке вище командування не знало, що бомби насправді падали в декількох милях від центру.⁴⁷

Очевидний план з'явився у британських військових командирів. Якби нацисти могли бути переконані, що їх бомби на вірній цілі, вони не змінили б свою траєкторію. А ще краще, якби вони були переконані, що бомби падають на північ від столиці, тоді вони скорегували б свою мету так, аби вони падали далі на південь – і, можливо, в кінцевому підсумку безпечно провалилися в сільській місцевості. Але ця змова означала, що бомби швидше приземляться на півдні Лондона.⁴⁸

Перехитрити ворогів вдалося – нацистам подавали неправдиву інформацію. Але відбулася жорстка дискусія між міністром з питань внутрішньої безпеки Гербертом Моррісоном та прем'єр-міністром Вінстоном Черчіллем. Було б занадто грубо називати це класовим конфліктом, але Моррісон, син поліцейського, можливо, відчував себе більш напружено, ніж Черчілль, ризикуючи людьми в районах робітничих класів південного Лондона. Черчілль, як завжди, переважав.⁴⁹

Серед істориків оскаржується успіх операції. Нацисти ніколи не вдосконалювали свою мету – і тим не менше, бомби забрали 6000 життів. Скільки ще могло бути втрачено, якби Черчілль не переважив у тій дискусії, можна лише здогадуватися.

Було б неправильно для Черчілля використовувати певних громадян як людський щит, навіть якщо його метою було б врятувати життя інших. Він би настільки ж помилявся, якби спрямовував людей на шлях нацистів, аби врятувати життя. Але, зрештою, він безумовно мав право підтримати змову,

⁴⁷ Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

⁴⁸ Там само.

⁴⁹ Там само.

яка перенаправила бомби на південь Лондона. В чому різниця? Допитливий випадок про товстуна на колії тримає ключ до відповіді.⁵⁰

Сценарії з вагонетками корисні саме тому, що вони настільки ретельно пропрацьовані. У звичайному житті повно сторонніх етичних шумів. Через складність реального життя важко виділити важливі якості морального міркування. Випадки з вагонетками вигадуються для того, аби виділити принципи та виявити важливі відмінності. Цього результату вони можуть досягти тільки в тому разі, якщо заглушать всі відволікаючі й спотворюючі шуми. Тут можна провести грубу аналогію з науковим методом. У лабораторії, якщо ви хочете, наприклад, вивчити дію світла, ви варіюєте параметри світла, підтримуючи всі інші фактори незмінними. Приблизно так само, якщо ви хочете з'ясувати, чи є морально значущим якийсь певний елемент, ви вигадуете два випадки, які тотожні у всьому, крім однієї змінної, з якою можна пограти.⁵¹

Як впливає з розгляду, дилема вагонетки і справді займає вагоме місце у моральній філософії та має беззаперечний позитивний вклад. У цьому розділі було розглянуто аналіз дилеми вагонетки сучасними філософами та її вплив не тільки на теперішній світ, а і на певні історичні події. За цими даними можна стверджувати, що дилема вагонетки відіграла певну позитивну роль в системі моральної філософії.

⁵⁰ Edmonds D. Matters of life and death / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>

⁵¹ Эдмондс Д. Убили бы вы толстяка? Задача о вагонетке: что такое хорошо и что такое плохо? / Д. Эдмондс [пер. с англ. Д.Кралечкина]. – М: Издательство Института Гайдара, 2016. – С. 136-137

РОЗДІЛ III

НЕГАТИВНИЙ АСПЕКТ ТА КРИТИКА ПРОБЛЕМИ

Мислинневі експерименти займали одне з провідних місць в багатьох роботах філософів. Деякі з них вносили їх як приклади у свої роботи, деякі давали їх на лекціях, аби зацікавити студентство. Та все ж, є люди, які вкрай негативно відгукуються про дилему вагонетки, а деякі вважають Філіппу Фут «тією, що створила монстра-Франкенштейна». В цьому я б і хотіла розібратися.

Робота Гріна та його колег над такими дилемами сильно вплинула на моральну філософію. Джошуа Грін – гарвардський психолог і нейровчений, що тяжів до утилітаризму. «Ось у чому справа», – подумав Грін. «Ось що відбувається в сценаріях з мостом і трансплантацією. Ми відчуваємо, що не повинні штовхати товстуна. Але думаємо, що краще врятувати життя п'ятьом, а не одному. Але думка і почуття – це різні речі».⁵² Отримавши одночасно психологічну та філософську освіту, Грін став першим, хто привернув увагу нейронауки до вивчення вагонеток. Він почав сканувати мозок людей в момент, коли їм надавали завдання з вагонетками. Сканери фіксували відповідні скачки в активності мозку. За описом Гріна, випадки з вагонетками провокують певну сутичку між ділянками мозку, що підраховують і емоційними ділянками. І в цій боротьбі у сторін набагато більш рівні шанси.⁵³

Коли пропонується дилема «товстуна» і вбивство, яке б ви зробили своїми власними руками, частини мозку, що знаходяться відразу за очима і вважаються відповідальними за такі почуття, як співчуття (мигдалеподібне тіло, кора задньої частини поясної звивини й префронтальна кора головного мозку), приходять в стан крайнього збудження. Думка про те, що треба штовхнути товстуна, включає у вас в мозку емоційний сигнал небезпеки, який говорить: «ні, це неправильно». Коли цього сигналу немає, ми включаємо

⁵² Цит. за Едмондс Д. Убили бы вы толстяка? Задача о вагонетке: что такое хорошо и что такое плохо? / Д. Эдмондс [пер. с англ. Д.Кралечкина]. – М: Издательство Института Гайдара, 2016. – С. 180

⁵³ Эдмондс Д. Убили бы вы толстяка? Задача о вагонетке: что такое хорошо и что такое плохо? / Д. Эдмондс [пер. с англ. Д.Кралечкина]. – М: Издательство Института Гайдара, 2016. – С. 180

звичайне утилітаристське числення: частина мозку, що підраховує (дорсолатеральна префронтальна кора і нижня тім'яна частка) та оцінює витрати й прибутки різних типів, не лише моральні. У стандартному випадку рівняння нескладне: ціною одного життя ми можемо врятувати п'ять. Вважається, що емоційні частини мозку розвинулися набагато раніше тих його ділянок, які відповідають за аналіз і планування. Тому в разі моральної дилеми можна очікувати на те, що емоція спровокує висновок швидше за розум. Дослідження показали, що, коли людей кваплять, вони проявляють менше схильності до утилітаризму.⁵⁴

Це призвело до більш ніж десятиліття захоплюючих досліджень з моральної психології та значних успіхів у нашому розумінні інструментальної шкоди. Однак, всупереч тому, наскільки популярні дилеми з жертвами були в моральній психології, цей підхід природно викликав певну критику, наприклад, стосовно високого штучного характеру зазвичай використовуваних сценаріїв. Моя мета тут – не надавати подальшої критики цієї парадигми, а виділити її межі як загальне мірило утилітарних тенденцій.⁵⁵

Дилеми з жертвами лише безпосередньо вимірюють те, що ми називаємо негативним виміром утилітаризму. Насправді, вони вимірюють лише ставлення до інструментальної шкоди – лише один аспект негативного виміру, хоч і дуже важливий, враховуючи моральну центральність заборон шкоди іншим. Судити про те, що, наприклад, нам слід штовхнути одну невинну людину з моста – це відхилити (або хоча б знизити) одне можливе деонтологічне правило проти прямого заподіяння шкоди комусь (як засіб запобігання більшої шкоди іншим). Але можна було б відкинути це особливе деонтологічне правило, все одно приймаючи багато інших правил – наприклад, правила, що стосуються справедливості, чесності або дотримання

⁵⁴ Едмондс Д. Убили бы вы толстяка? Задача о вагонетке: что такое хорошо и что такое плохо? / Д. Эдмондс [пер. с англ. Д.Кралечкина]. – М: Издательство Института Гайдара, 2016. – С. 182

⁵⁵ Kahane G., Everett J., Earp B. та ін. Beyond sacrificial harm: A two-dimensional model of utilitarian psychology / Psychological Review. – 2018. – №125. – P. 136.

обіцянок. І, безумовно, це правило можна відкинути, залишаючись дуже частковим у прийнятті моральних рішень.⁵⁶

Якщо хтось ухвалює рішення, яке відповідає конкретному утилітаризму, це не відразу свідчить про те, що ці судження випливають із міркувань, що лежать в основі утилітарного морального світогляду, або відповідають йому. Це також не свідчить, що хтось буде судити за таким світоглядом й в інших контекстах. Однак, існує емпірична можливість, що відповіді на моральні дилеми відображають застосування загального утилітарного принципу (неупереджено максимізувати очікувану корисність) або навіть вказують на більш широку тенденцію підходу до прийняття моральних рішень утилітарним способом. Якби це було так, це потенційно могло б виправдати майже ексклюзивну спрямованість на дилему, оскільки утилітарне прийняття рішень було б унітарним психологічним явищем у контексті непрофесійного населення, із загальним набором процесів та факторів, що призвели б до суджень згідно з утилітаризмом у ряді моральних контекстів. Вивчаючи відповіді на дилему вагонетки, ми б висвітлювати це загальне явище. Це припущення є імпліцитним у дослідженнях Гріна, які в основному зосереджені на дилемах вагонетки, але все ж мають загальні претензії щодо утилітарного прийняття рішень та суджень. Але припущення, що утилітарне прийняття рішень – це цілісний феномен, який повністю (або навіть унікально) зображається у реакціях на дилему вагонетки, також, здається, лежить в основі багатьох інших досліджень, що застосовують такі дилеми, згідно з якими результати регулярно зазначаються як підтримка загальних висновків щодо психології утилітарного прийняття рішень. І все ж, асоціація між «утилітарними» судженнями щодо дилем та антисоціальними рисами, такими як психопатія як клінічної, так і субклінічної популяції ставлять під сумнів це припущення. Інші дослідження показали взаємозв'язок між судженнями про жертвопринесення і політичними поглядами на свободу, а також явним

⁵⁶ Kahane G., Everett J., Earp B. та ін. Beyond sacrificial harm: A two-dimensional model of utilitarian psychology / Psychological Review. – 2018. – №125. – P. 136-137.

схваленням різних форм егоїзму, обидва з яких надають підвищений або навіть винятковий пріоритет власному інтересу над добробутом інших.⁵⁷

Психопати, очевидно, не є зразком неупередженого занепокоєння про благо, і егоїсти явно відкидають будь-яке подібне занепокоєння. Такі висновки прямо суперечать суворій неупередженості людських інтересів, які лежать в основі утилітарної теорії. Подальший відхід від такої неупередженості вбачається в дослідженнях, що дозволяють припустити – рішення про скоєння зла є більш імовірними, якщо вони є в інтересах учасників; а також і в дослідженнях, що дозволяють припустити факт, що частота рішення про скоєння зла знаходиться під впливом того, чи людина перебуває в групі; чи є порівняння між іноземцями та земляками, незнайомцями чи членами сім'ї, або навіть тваринами чи людьми. Такі висновки говорять про те, що судження про скоєння зла у цих дилемах рідко базуються на такій собі неупередженій максимізації сукупного добробуту, якого вимагає утилітаризм.⁵⁸

За утилітаризмом, ми маємо максимально збільшувати загальний добробут кожного, і маємо право серйозно нашкодити або навіть вбити невинних людей, якщо це потрібно для досягнення всеохоплюючого морального ідеалу. Більшість психологічних досліджень утилітарного прийняття рішень до цього часу було зосереджено саме на цьому останньому, негативному вимірі утилітаризму, значною мірою ігноруючи більш фундаментальний неупереджений ідеал, який лежить в його основі. Шкала Оксфордського утилітаризму (OUS) націлена на усунення обмежень попередніх заходів, особливо стандартного в цей час методу вимірювання так званих «утилітарних» суджень жертвоприношеннями. Висновки показують, що протоутилітарні тенденції не утворюють унітарного психологічного явища у суспільстві, а насправді складаються з двох значною мірою незалежних компонентів. Це пояснює дивовижні результати в чинній літературі, відкриває

⁵⁷ Kahane G., Everett J., Earp B. та ін. Beyond sacrificial harm: A two-dimensional model of utilitarian psychology / Psychological Review. – 2018. – №125. – P. 136-137.

⁵⁸ Там само. P. 155.

для нас психологічні джерела як утилітаризму, так і широкого протистояння йому, і висвітлює нові напрямки для майбутніх досліджень.⁵⁹

Утилітаризм часто зображається його прихильниками як продукт чітко вираженої раціональної рефлексії та стійкості, що значною мірою пояснюється тим, як він стикається з потужними інтуїціями та емоціями (такими, які викликані ідеєю насильницької жертви невинної людини). Грін стверджував, що замість того, аби зрозуміти деонтологію та консеквенціалізм як «філософські винаходи», їх краще розуміти як «філософські прояви двох нерозбірливих психологічних зразків, двох різних способів морального мислення, які були частиною людського репертуару на тисячі років». Підтримка утилітарних (консеквенціоністських) або деонтологічних етичних принципів майже напевно визначається психологічними диспозиціями та процесами, які є частиною нашого основного ментального інструментарію; ми не можемо погодитись із думкою, що утилітарна сторона рівняння є результатом єдиної психологічної закономірності. Ці висновки говорять про те, що в утилітарному вимірі є два нерозбірливі психологічні джерела: одне, що стосується неупередженої користі та інше щодо інструментальної шкоди. Іншими словами, для утилітарного прийняття рішень є два невід'ємні компоненти – щось зовсім відмінне від моделі подвійного процесу Гріна, що зображає дві системи, які лежать в основі деонтології проти утилітаризму, при цьому трактуючи кожен як єдину психологічну конструкцію. У моделі Гріна це «дуель» автоматичних проти контрольованих процесів, що лежать в основі відповідно деонтологічних проти утилітарних суджень; в нашій моделі ці дві сторони відносяться до двох різних компонентів утилітаризму.⁶⁰

Незважаючи на те, що унітарна модель утилітаризму як когнітивна (і деонтологія як емоційна), здавалося б, підтримується більш ранніми дослідженнями, більш сучасна робота пов'язана з подібними судженнями

⁵⁹ Kahane G., Everett J., Earp B. та ін. Beyond sacrificial harm: A two-dimensional model of utilitarian psychology / Psychological Review. – 2018. – №125. – P. 155.

⁶⁰ Там само. P. 158.

щодо зменшення відрази до шкоди. Переконання про жертвопринесення значною мірою спричинені зменшенням емоцій. Позитивний вимір утилітаризму також часто стверджується на раціональній рефлексії – відомий аргумент Пітера Сінгера про те, що немає моральної різниці між дозволом дитині потонути та відмовою до пожертвування грошей на країни, що розвиваються; аргумент ґрунтується на заклику до послідовності, а не на гру з емоціями. Однак, це дослідження також пов'язує неупереджену користь із суперечливим занепокоєнням, афективним настроєм – дійсно зворотним до того, що пов'язане з інструментальною шкодою. Ці результати узгоджуються з останніми дослідженнями, які виявили, що альтруїсти, які пожертвували свої нирки незнайомцям, виявляють більшу емпатичну стурбованість.⁶¹

З доктрини утилітаризму та її застосування випливають певні наслідки, які багато хто вважає дуже парадоксальними й навіть огидними. Люди, які тяжіють як до інструментальної шкоди, так і до неупередженої користі, швидше за все, будуть схилитися до утилітаризму. Однак психологічна незалежність та ступінь напруження між цими двома вимірами утилітаризму свідчать про подальшу перешкоду для його прийняття: такі психологічні фактори як емпатична стурбованість, можуть зробити деяких людей сприйнятливими до одного виміру, та стійкими до іншого. Можливо, існує певна психологічна нестабільність, що стосується утилітаризму. Як зазначалося вище, отримані результати свідчать про необхідні подальші дослідження для з'ясування причинно-наслідкового зв'язку між співпереживанням емпатії та неупередженої користі.⁶²

На підставі сказаного вище можна дійти висновку, що дилема вагонетки має свої негативні аспекти, такі як двобічність утилітаризму; з однієї сторони ставлення з неупередженою користю задля благополуччя всіх людей, а з іншої – холодне, беземоційне сприйняття інструментальної шкоди. Аналіз показав,

⁶¹ Kahane G., Everett J., Earp B. та ін. Beyond sacrificial harm: A two-dimensional model of utilitarian psychology / Psychological Review. – 2018. – №125. – P. 158.

⁶² Там само. С.159.

що обидва варіанти мають свої вагомі мінуси, а аспект утилітаризму все ще потребує подальших наукових досліджень.

РОЗДІЛ IV

МОРАЛЬНИЙ ВИБІР У ШТУЧНОГО ІНТЕЛЕКТУ

Сучасні революційні досягнення науково-технічного прогресу поставили важливе питання: чи можуть мораль і право сприйматися інтелектом, що є розбіжним з інтелектом людини; та й чи може інтелект керуватися мораллю і правом у своїй діяльності. Людський інтелект частіше за все описують так: здібність до пізнання навколишнього світу, критичне мислення, можливість діяти в межах знакової системи й самостійне прийняття рішень. Але, штучний інтелект найвищого ступеня розвитку вже визнаний більш розумним, ніж інтелект будь-якої людини практично в усіх галузях; він володіє навіть більш потужними когнітивними властивостями, такими як обробка великих обсягів інформації, нейтральна об'єктивна оцінка ситуації, стратегічність та тактичність форми поведінки, планування та переорієнтація в залежності від ситуації.

Жоден об'єкт техніки (доки не з'явився штучний інтелект) не міг порівнюватись з людиною стосовно її здатності до самоконтролю, підтримки життєдіяльності та пристосування до умов навколишнього середовища. Але наразі штучний інтелект вже переважає людину розумінням своєї власної внутрішньої будови та здатністю до перебудови задля подальшого нескінченного вдосконалення (прискорена еволюція).⁶³

Неважко передбачити, що так звана цифрова людина підведеться як мінімум на одну сходинку вище в умовній еволюційній ланці у порівнянні зі звичайною людиною. Людина цифрової доби, що буде вдосконалюватись за певним напрямком, може дорівнювати можливостям штучного інтелекту та помітно відрізнитися від інтелекту звичайної людини. Але чи буде вона

⁶³ Радутний О.Е. Мораль і право для штучного інтелекту та цифрової людини: закони робототехніки та «проблема вагонетки» / Інформація і право: науковий журнал / редкол.: В.Г. Пилипчук та ін. – К.: Науково-дослідний інститут інформатики і права Національної академії правових наук України, 2019. – № 3 (30). – с.81.

визнавати такі ж моральні цінності та дотримуватися таких самих поглядів на права та свободи звичайної людини?⁶⁴

Втім, основною метою появи штучного інтелекту та самостійних транспортних засобів є замінювання ними людини у небезпечних ризикових сферах діяльності, бажання зменшити аварійність, запобігання помилкам, які мають наслідки у людському факторі. Відтак, згідно з виданням *The Economist*, 90% аварій на дорогах відбувається через помилки людини. Тому надати право на керування того ж транспорту штучному інтелектові має велику кількість переваг. Скажімо, якщо говорити про автономні транспортні засоби, є певне сподівання, що на дорогах будуть відсутні порушення правил дорожнього руху, буде присутня швидка орієнтація в складній ситуації, буде відсутнє емоційне забарвлення/почуття страху, буде врахована значна кількість супутніх даних: дії інших учасників руху, робота світлофорів, камер спостереження по всьому маршруту, завантаженість окремих транспортних шляхів тощо. Штучний інтелект безумовно буде більш законослухняним та буде діяти передбачувано у порівнянні з поведінкою звичайної людини; все ж, як нам повідомляє *The Guardian*, перші автомобілі з автопілотом (Volvo) з'являться на лондонських дорогах без жодних помітних знаків, що має на меті попередити можливу зміну поведінки інших учасників дорожньо-транспортного руху, оскільки велика кількість опитаних респондентів заявили, що схилиються до розглядання автономного транспортного засобу як простого об'єкта для взаємодії, ніж авто, якими керує людина; адже вони так чи інакше будуть завжди дотримуватись всіх вимог та правил безпеки руху, що може підштовхнути певних водіїв (фізичних осіб) до зловживань.⁶⁵

Дослідники Колумбійського університету (Columbia University) вважають, що використання автономних транспортних засобів незабаром

⁶⁴ Радутний О.Е. Мораль і право для штучного інтелекту та цифрової людини: закони робототехніки та «проблема вагонетки» / Інформація і право: науковий журнал / редкол.: В.Г. Пилипчук та ін. – К.: Науково-дослідний інститут інформатики і права Національної академії правових наук України, 2019. – № 3 (30). – с.82.

⁶⁵ Там само. С. 83.

збільшить пропускну здатність доріг на 273%, а середня швидкість руху на магістралях зросте вдвічі. Як прогнозує Інститут інженерів з електротехніки та електроніки (Institute of Electrical and Electronics Engineers) до 2045 року майже 75% транспортних засобів перейдуть на автономний режим. Втім, більшість опитаних людей проголосували за придбання (або користування) автономним транспортним засобом, що в певній ситуації може захистити саме їх на протипагу іншим учасникам дорожнього руху. Гіпотетично, вченими пропонується спрогнозувати перспективу підключення штучного інтелекту (для керування транспортним засобом) до соціальних мереж задля розуміння того, хто піддаватиметься ризику. Коли більшість людей замислюються про керування або їзду в автономному транспортному засобі, стрімко зменшується цікавість щодо вирішення моральних дилем, а вибір лежить скоріше на користь власного виживання. Однак, в певних випадках неспіврозмірної шкоди може порушуватись баланс крайньої необхідності – рішення буде прийматись штучним інтелектом, а не людиною. Проте, не виключено, що варіант (наприклад рятувати пасажирів автономного транспортного засобу, але з дотриманням правил дорожнього руху; чи жертвувати одним пасажиром заради врятування дитини-пішохода за умови дотримання правил) можна буде активувати ще на початку використання транспортного засобу з автопілотом.⁶⁶

Окрім цього, не важко уявити сценарій, коли алгоритми автономного керування можуть увійти у конфлікт між собою. Однією з перших спроб подолати таку проблему і надати штучному інтелектові більш точний навігатор була у Німеччині у 2016 році. За інформацією видання Inverse, міністр з транспорту в Німеччині Олександр Добріндт (Alexander Dobrindt) висунув щонайменш три етичні правила, яких мають дотримуватися будь-який автономний транспорт під час керування:

⁶⁶ Радутний О.Е. Мораль і право для штучного інтелекту та цифрової людини: закони робототехніки та «проблема вагонетки» / Інформація і право: науковий журнал / редкол.: В.Г. Пилипчук та ін. – К.: Науково-дослідний інститут інформатики і права Національної академії правових наук України, 2019. – № 3 (30). – с. 84.

- 1) Пошкодження майна у будь-якому випадку краще, ніж нанесення шкоди людині;
- 2) Заборонено класифікувати людей за розміром, віком або іншими параметрами;
- 3) У будь-якому випадку, відповідальність лягає на виробника.

Окрім того, міністр висловився в інтерв'ю WirtschaftsWoche про обладнання кожного авто з автопілотом «чорною скринькою» та обов'язкову присутність живого водія у ньому.⁶⁷

Правило «пошкодження майна у будь-якому випадку краще, ніж нанесення шкоди людині» повністю відповідає вимогам крайньої необхідності (ст.39 КК України). Друге правило «заборонено класифікувати людей за розміром, віком або іншими параметрами» є співзвучним до положень Конвенції про захист прав людини і основоположних свобод та іншим сучасним системоутворюючим документам міжнародного рівня. Але третє правило «у будь-якому випадку, відповідальність лягає на виробника» входить у суперечність, принципами правової держави та правової визначеності. Штучному інтелекту лишається лише дотримуватись норм права, що є публічними та загальнообов'язковими. Безсумнівно, з цим завданням штучному інтелектові буде дуже легко впоратись, у порівнянні з середньостатистичною людиною. Проте, більш цікавим виглядає питання свідомої відсутності у штучного інтелекту прихильності до якогось аспекту моралі, зокрема, крайніх чи агресивних поглядів.⁶⁸

Норми моралі людина осягає упродовж всього життя, відтак результат базових знань слід демонструвати лише на окремо визначених етапах соціальної зрілості (наприклад, перехід від стану хлопчика до стану чоловіка, присяга військового, ініціація під час шлюбу, що означає вмирання старого і

⁶⁷ Радутний О.Е. Мораль і право для штучного інтелекту та цифрової людини: закони робототехніки та «проблема вагонетки» / Інформація і право: науковий журнал / редкол.: В.Г. Пилипчук та ін. – К.: Науково-дослідний інститут інформатики і права Національної академії правових наук України, 2019. – № 3 (30). – с. 85.

⁶⁸ Там само. С. 86.

народження нового статусу тощо). Для штучного інтелекту не виявиться складним засвоєння, аналіз та систематизація відповідних моральних та правових приписів у більш короткий проміжок часу, аніж потрібно звичайній людині (секунди/години замість десятків років). Отже, на принципове питання про те, чи засвоїть штучний інтелект норми моралі, відповідь – скоріше за все, так; штучний інтелект у певних випадках може бути навіть більш моральним, ніж людина.⁶⁹

Ніколи в історії людства ми не дозволяли машині самостійно вирішувати, хто повинен жити, а хто повинен померти, за секунду, без огляду в реальному часі. Ми збираємось перетнути цей міст будь-коли, і це не станеться у віддаленому театрі військових дій; це відбудеться в тому, що є найназемнішим аспектом нашого життя – щоденним транспортом. Перш ніж дозволити нашим автомобілям приймати етичні рішення, нам потрібно провести глобальну дискусію, аби висловити свої переваги компаніям, які розробляють моральні алгоритми, та політикам, які їх регулюватимуть. Для початку такої розмови було розроблено Моральну машину (the Moral Machine – онлайн-експериментальна платформа, призначена для вивчення моральних дилем, з якими стикаються автономні транспортні засоби), і мільйони людей з усього світу взяли участь. Респонденти могли б бути такими ж скупими чи ретельними, як хотіли б у етичних рамках, яких вони вирішили дотримуватися. Вони могли брати участь у складному зважуванні всіх дев'яти рішень, що використовуються в Моральній машині, або прийняти прості правила, такі як «нехай машина завжди їде вперед». Ці дані допомогли визначити три сильні вподобання, які можуть слугувати складовими для дискусій про загальну етику машин, навіть якщо вони в остаточному підсумку не схвалюються політиками: уподобання до рятування людського життя, перевага до рятування більшої кількості життів та перевага до рятування молодших. Деякі

⁶⁹ Радутний О.Е. Мораль і право для штучного інтелекту та цифрової людини: закони робототехніки та «проблема вагонетки» / Інформація і право: науковий журнал / редкол.: В.Г. Пилипчук та ін. – К.: Науково-дослідний інститут інформатики і права Національної академії правових наук України, 2019. – № 3 (30). – с. 87.

вподобання, що ґрунтуються на гендерному та соціальному статусі, значно відрізняються в різних країнах і, схоже, відображають основні переваги суспільного рівня щодо егалітаризму.⁷⁰

Проект «Моральна машина» багато в чому був нетиповим. Нетиповим у своїх цілях та амбіціях: жодне дослідження раніше не намагалося виміряти моральні вподобання, використовуючи дев'ятифакторний експериментальний дизайн у більш ніж 200 країнах. Для досягнення цієї незвичної мети було розгорнуто вірусну онлайн-платформу, сподіваючись охопити величезну кількість учасників. Це дозволило збирати дані від мільйонів людей у всьому світі, подвиг, який було б майже неможливо досягти за допомогою стандартних методів академічного опитування. Наприклад, набір національно репрезентативних зразків учасників у сотень країн уже був би вкрай складним, але перевірити дев'ятифакторний дизайн у кожному з цих зразків виявилось б неможливим. Цей підхід дозволив обійти ці складнощі, але його недоліком є те, що зразок обирається самостійно і не гарантовано буде відповідати соціально-демографічній ситуації кожної країни. Той факт, що міжсоціальна зміна збігається з раніше встановленими культурними кластерами, а також той факт, що макроекономічні змінні прогнозують відповіді «Моральної машини», є хорошими сигналами. Але той факт, що зразки не гарантовано є репрезентативними, означає, що політики не повинні сприймати ці дані як остаточне слово щодо суспільних уподобань – навіть якщо цей зразок, можливо, близький до технічно-підкованого населення, яке зацікавлене в автопілоті машин.⁷¹

Навіть при величині вибірки, не можливо зробити справедливою всю складність автономних дилем. Наприклад, цих сценаріях персонажів визнавали дорослими, дітьми, тощо з 100% впевненістю, а результати життя та смерті прогнозувалися зі 100% впевненістю. Ці припущення є технологічно нереальними, але вони були необхідні для того, аби проект можна було

⁷⁰ Awad E. The Moral Machine experiment / E. Awad, S. Dsouza, R. Kim. // Nature. – 2018. – P. 63.

⁷¹ Там само. С. 63.

відстежувати та дослідити. Так само не маніпулювалося гіпотетичними відносинами між респондентами та персонажами (наприклад, родичі або подружжя).⁷²

Дійсно, ми можемо сприймати виклики машинної етики як унікальну можливість вирішити, як громада, що ми вважаємо правильним чи неправильним; і переконатися, що машини, на відміну від людей, безпомилково дотримуються цих моральних уподобань. Ми не можемо досягти універсальної згоди: навіть найсильніші вподобання, висловлені через «Моральну машину», показали істотні культурні розбіжності, і цей проект спирається на давню традицію дослідження культурних варіацій етичних суджень. Але той факт, що різні регіони світу виявили певну згоду, говорить про те, що наша подорож до консенсусної машинної етики не була приречена з самого початку. Часто рекомендують, аби етика машин повинна бути узгоджена з людськими цінностями. Водночас, люди можуть відчувати внутрішній конфлікт, міжособистісні розбіжності та культурні відмінності в моральній сфері. В цьому експерименті показано, що ці конфлікти, розбіжності та відмінності, хоча й істотні, не можуть бути фатальними.⁷³

Як впливає з розгляду, моральні дилеми є особливою актуальною проблемою в контексті штучного інтелекту (а саме автономних транспортних засобів). Проаналізувавши теперішній стан щодо питання морального вибору у машинній етиці, я дійшла висновку, що ця проблема ще не досягла своєї кульмінації – наразі результати є доволі нейтральними. З однієї сторони, ми можемо запрограмувати штучний інтелект і ніби-то «навчити» його вирішувати моральні дилеми; з іншої – етика машин ніколи не зможе бути до кінця узгодженою з людськими цінностями, відчувати внутрішні конфлікти, міжособистісні та культурні розбіжності – і враховувати їх при цьому надважливому виборі.

⁷² Awad E. The Moral Machine experiment / E. Awad, S. Dsouza, R. Kim. // Nature. – 2018. – P. 63.

⁷³ Там само. С. 63.

ВИСНОВКИ

Нарешті, існує багато версій дилеми вагонетки, кількість яких зростає з кожним днем у різних сферах та галузях. За співвідношенням дилеми вагонетки із теоріями справедливості, можна зазначити про кантіанську, утилітаристську концепції та доктрину подвійного ефекту. Значення дилеми вагонетки полягає у неоцінімому внеску не тільки у філософський напрямок, а й у психологію, нейретику та когнітивну науку. А це означає, що дана проблема все ще залишається відкритою та буде розвиватися впродовж багатьох століть.⁷⁴

За останні півстоліття вагонеткологія виступала інструментом, що дозволяє оскаржити фундаментальні положення етики і розглянути життєво важливі питання про те, як ми повинні ставитися один до одного і жити своїм життям. Коли Філіпа Фут вигала завдання з вагонеткою, вона була орієнтована на суперечку з приводу абортів. Сьогодні ж завдання, що нагадують сценарії з вагонетками, частіше виникають в обговореннях легітимності різних типів військових дій. Дилема Черчілля – чи варто намагатися перенаправити ракети в менш населені райони – зустрічається сьогодні в нових втіленнях. Завдання з «товстуном» відтіняє непримиренне зіткнення деонтологічної й утилітаристської етики. У більшості людей немає утилітаристських інстинктів (що визнають і самі утилітаристи). Вони вважають, що, можливо, Вінстон Черчілль вчинив погано, коли використав громадян як живий щит, навіть якщо його мета полягала в тому, аби врятувати життя іншим. Але в цілому він напевне вчинив правильно, підтримавши план з дезінформації та відведення ракет до південного Лондону. Чому існує ця різниця? Філософи все ще не можуть дійти до єдиної думки. Але що б вони не відповіли, ключем до відповіді має бути дивне положення товстуна на пішохідному мосту.⁷⁵

⁷⁴ Сатохіна Н. І. «Проблема вагонетки» і сучасні теорії справедливості / Н. І. Сатохіна, Л. А. Улашкевич. // Молодий вчений. – 2018. – С. 3.

⁷⁵ Эдмондс Д. Убили бы вы толстяка? Задача о вагонетке: что такое хорошо и что такое плохо? / Д. Эдмондс [пер. с англ. Д.Кралечкина]. – М: Издательство Института Гайдара, 2016. – С. 136

Звісно, дилема вагонетки має свої евристичні переваги і недоліки. Багато дослідників дякують Філіппі Фут за її існування, і, завдяки низці наукових доробок, що ми маємо на даний момент, розуміння моральної етики може позитивно впливати на втілення таких інновацій як автопілот машин та в майбутньому й інших пристроїв. Так само, як позитивний вплив, проблема має і негативний вплив та критику: одні пишуть, що як заходи утилітарних рішень, жертвні дилеми, такі як дилема вагонетки, вимірюють лише одну грань протоутилітарних тенденцій, а саме вседозволеного ставлення до інструментальної шкоди, ігноруючи при цьому неупереджену стурбованість більшою користю; дилема передбачає лише часткову міру утилітаризму. Інші критикують, що сценарій, який представляє дилема, є надто екстремальним і не пов'язаним з моральними ситуаціями в реальному житті.⁷⁶

У своїй роботі я розглянула суперечність морального вибору в дилемі вагонетки. Були виконані всі завдання, що були поставлені у цій роботі: я розглянула історію проблеми та її різноманітні варіанти у першому розділі; були проаналізовані як першопочатковий варіант дилеми, запропонований Філіппою Фут, так і різноманітні сучасні варіації, що супроводжувались соціологічними дослідженнями. Також був досліджений аспект суперечності деонтології (а саме кантіанської етики) та утилітаризму в моральній дилемі. У другому розділі я проаналізувала позитивну роль дилеми в системі моральної філософії; були розглянуті дослідження сучасних філософів щодо дилеми та вплив дилеми не тільки на теперішній світ, а і на певні історичні події. У третьому розділі я спробувала зосередитись на негативному аспекті дилеми вагонетки та критиці цієї проблеми, а саме на двобічності утилітаризму та суперечність неупередженої користі та інструментальної шкоди. У четвертому розділі я розглянула актуальне питання морального вибору в майбутньому та моральну етику в контексті штучного інтелекту; аналіз в основному

⁷⁶ Bauman C. Revisiting External Validity: Concerns about Trolley Problems and Other Sacrificial Dilemmas in Moral Psychology / C. Bauman, P. McGraw, D. Bartels. // *Social and Personality Psychology Compass*. – 2014. – №8. – С. 546.

зосереджувався на автономних транспортних засобах. Ця проблема не є дослідженою до кінця та наразі її результати є дещо нейтральними. З однієї сторони, штучний інтелект може бути запрограмований і «навчений» вирішувати моральні дилеми; а з іншої – етика машин ніколи не зможе бути до кінця узгодженою з людськими цінностями, оскільки важливо відчувати міжособистісні та культурні розбіжності, і враховувати їх при цьому виборі. На мою думку, суперечності в сприйнятті цієї дилеми будуть жити ще довго, важливо не тільки знати про ці суперечності, а й аналізувати їх для майбутніх досліджень.

Що ж буде з майбутнім? Чи зможе штучний інтелект обрати, чи повернути перемикач або ж чи зіштовхнути товстуна? На цій стадії все залежить від нас. Ми – те саме покоління, яке буде програмувати і систематизувати штучний інтелект, налаштовувати його на свої моральні вподобання і етичні правила. Звісно, як вже було сказано раніше, машина не буде відчувати внутрішнього конфлікту, міжособистісних розбіжностей чи культурних відмінностей, але, як на мене, в наших силах навчити машину обирати. Ті наукові доробки, що десятиліттями складали пазл моральної етики стануть впевненим підґрунтям для розвитку машинної етики, і ми можемо впевнено казати, що дилема вагонетки існує не дарма.

Дилема вагонетки і справді є одним з найбільш суперечливих в етиці моральних виборів. В одному випадку мені захочеться, аби не постраждала невинна людина, в іншому – захочеться врятувати більшість. В деяких випадках однозначним переможцем буде деонтологія, в усіх інших її перемагатиме утилітаризм. Чи має існувати «правильний» чи «неправильний» вибір? На це питання в мене немає відповіді. Але я точно знаю, що на питання «чи не даремно існує дилема вагонетки?» відповідь буде позитивною.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Архелюк В. В. Актуальні проблеми структурної лінгвістики / В. В. Архелюк. // Молодий вчений. – 2015. – С. 175–179.
2. Радутний О.Е. Мораль і право для штучного інтелекту та цифрової людини: закони робототехніки та «проблема вагонетки» / Інформація і право: науковий журнал – К.: Науково-дослідний інститут інформатики і права Національної академії правових наук України, 2019. – № 3 (30). – 142 с. – с. 78 – 96
3. Сатохіна Н. І. «Проблема вагонетки» і сучасні теорії справедливості / Н. І. Сатохіна, Л. А. Улашкевич. // Молодий вчений. – 2018. – С. 1–4.
4. Скворцов А. А. Дж. Дж. Томсон: убийство, позволение умереть и проблема трамвая / А. А. Скворцов. // Философия и общество. – 2019. – С. 135–157.
5. Эдмондс Д. Убили бы вы толстяка? Задача о вагонетке: что такое хорошо и что такое плохо? / Д. Эдмондс [пер. с англ. Д.Кралечкина]. – М: Издательство Института Гайдара, 2016. – 256 с.
6. Awad E. The Moral Machine experiment / E. Awad, S. Dsouza, R. Kim. // Nature. – 2018. – P. 59–64.
7. Bauman C. Revisiting External Validity: Concerns about Trolley Problems and Other Sacrificial Dilemmas in Moral Psychology / C. Bauman, P. McGraw, D. Bartels. // Social and Personality Psychology Compass. – 2014. – №8. – P. 536–554.
8. Bleske-Rechek A. Evolution and the trolley problem: People save five over one unless the one is young, genetically related, or a romantic partner. / A. Bleske-Rechek, L. Nelson, J. Baker. // Journal of Social, Evolutionary, and Cultural Psychology. – 2010. – №4. – P. 115–127.
9. Bruers S. A review and systematization of the Trolley Problem [Електронний ресурс] / S. Bruers, J. Braeckman // Springer Science. – 2013. – Режим доступу: <https://rdcu.be/b345Z>

10. Cathcart T. The Trolley Problem, or Would You Throw the Fat Guy Off the Bridge?: A Philosophical Conundrum / Thomas Cathcart. – N. Y.: Workman Publishing Company, Inc., 2013. – 125 p.
11. Consequentialism [Електронний ресурс] // Stanford Encyclopedia of Philosophy. – 2019. – Режим доступу: <https://plato.stanford.edu/entries/consequentialism/>
12. Chomsky N. The 'Chomskyan Era' / Noam Chomsky // The Architecture of Language. – 2000. – Режим доступу до ресурсу: <https://chomsky.info/architecture01/>
13. Edmonds D. Matters of life and death [Електронний ресурс] / David Edmonds // Prospect. – 2010. – Режим доступу: <https://www.prospectmagazine.co.uk/magazine/ethics-trolley-problem>
14. Hauser M. Reviving Rawls' linguistic analogy: operative principles and the causal structure of moral actions [Електронний ресурс] / M. Hauser, L. Young, F. Cushman // Moral Psychology and Biology, Oxford U. Press, NY. – 2008. – Режим доступу: https://www.academia.edu/2050790/Reviving_Rawls_linguistic_analogy
15. Foot P. The Problem of Abortion and the Doctrine of the Double Effect / Philippa Foot // Oxford Review, No. 5. [Електронний ресурс] – 1967. – Режим доступу: <https://philpapers.org/archive/footpo-2.pdf>
16. Kahane G., Everett J., Earp B. та ін. Beyond sacrificial harm: A two-dimensional model of utilitarian psychology / Psychological Review. – 2018. – №125. – P. 131–164.
17. Mind Field (2017) The Greater Good [освітнє відео] // YouTube. 1 лютого. Режим доступу: <https://youtu.be/1sl5KJ69qiA>
18. Navarrete C. Virtual morality: Emotion and action in a simulated three-dimensional "trolley problem" / C. Navarrete, M. McDonald, M. Mott. // Emotion. – 2012. – №12. – С. 364–370.
19. Otsuka M. Kamm on the morality of killing / Michael Otsuka. // Ethics. – 1997. – С. 197–207.

20. Protocol additional to the Geneva Conventions of 12 August 1949, and relating to the protection of victims of international armed conflicts (Protocol I) (adopted 8 June 1977, entered into force 7 December 1978). Режим доступа:
<https://treaties.un.org/doc/Publication/UNTS/Volume%201125/volume-1125-I-17512-English.pdf>
21. Thomson J. The Trolley Problem / Judith Jarvis Thomson. // Yale Law Journal. – 1985. – №94. – P. 1395–1415.
22. Trammell R. Saving Life and Taking Life / Richard Trammell. // Journal of Philosophy. – 1975. – №72. – С. 131–137.
23. Unger P. Living high and letting die / Peter Unger. – Oxford: Oxford University Press, 1996. – 200 p.
24. Waldmann M. Throwing a bomb on a person versus throwing a person on a bomb. Intervention myopia in moral intuitions [Электронный ресурс] / M. Waldmann, J. Dieterich // Psychological Science. – 2007. – Режим доступа:
https://www.psych.uni-goettingen.de/de/cognition/publikationen-dateien/waldmann/2007_bomb.pdf