



Как доказать, что другие люди что-то чувствуют

Ты знаешь о своей боли изнутри, а о чужой — только по внешним признакам. Разница больше, чем кажется



Философия · 13–17 лет · 15 минут

Наблюдать можно поведение: крик, слёзы, слова. Само переживание наблюдать нельзя ни у кого, кроме себя. Из этого простого разрыва растут вопросы о животных, о машинах и о том, почему мы вообще считаем других людей живыми.

Разрыв, который трудно заметить

Ты знаешь, что такое боль. Знаешь изнутри, непосредственно: когда болит, никакое доказательство не нужно.

Теперь про другого человека. Он морщится, отдёргивает руку, говорит «больно». Из этого ты заключаешь, что ему больно.

Но заметь, что именно ты наблюдал. Ты наблюдал **поведение**: движение мышц, звук, слова. Само переживание ты не наблюдал и наблюдать не можешь. Ни в каком микроскопе нельзя увидеть чужую боль — можно увидеть только тело, которое ведёт себя как болящее.

Отсюда вопрос, который выглядит безумным и при этом строго осмыслен: **на каком основании ты считаешь, что за этим поведением вообще что-то есть?**

Может быть, все остальные устроены так, что ведут себя совершенно как люди, говорят о своих чувствах, плачут и смеются — а внутри у них не происходит ничего. Темнота. Никакого переживания.

Проверить нельзя. Любая проверка даст тебе ещё поведение.

Почему это не пустая игра

Сначала защитимся от возражения «да очевидно же, что они чувствуют».

Вывод действительно правильный — почти наверняка все чувствуют. Интересно не это, а **чем именно он обоснован**, потому что обоснование оказывается слабее, чем кажется, и от этого зависят вполне практические вещи.

Обоснование обычно такое: другие устроены как я, ведут себя как я, значит, внутри у них то же, что у меня.

Это называется рассуждением по аналогии, и у него есть неприятная особенность: **оно опирается на единственный случай**. Ты знаешь изнутри ровно одно существо — себя. Делать общий вывод из выборки в одного — в любой другой области недопустимо.

Работает такой довод ровно постольку, поскольку другой похож на тебя. И вот тут начинается самое интересное: **чем меньше сходства, тем слабее вывод**. Для другого человека он очень силён. Для со-



Наблюдаемо только внешнее, переживается только своё

баки слабее. Для осьминога ещё слабее. Для насекомого совсем слабый. Для программы — непонятно, применим ли вообще.

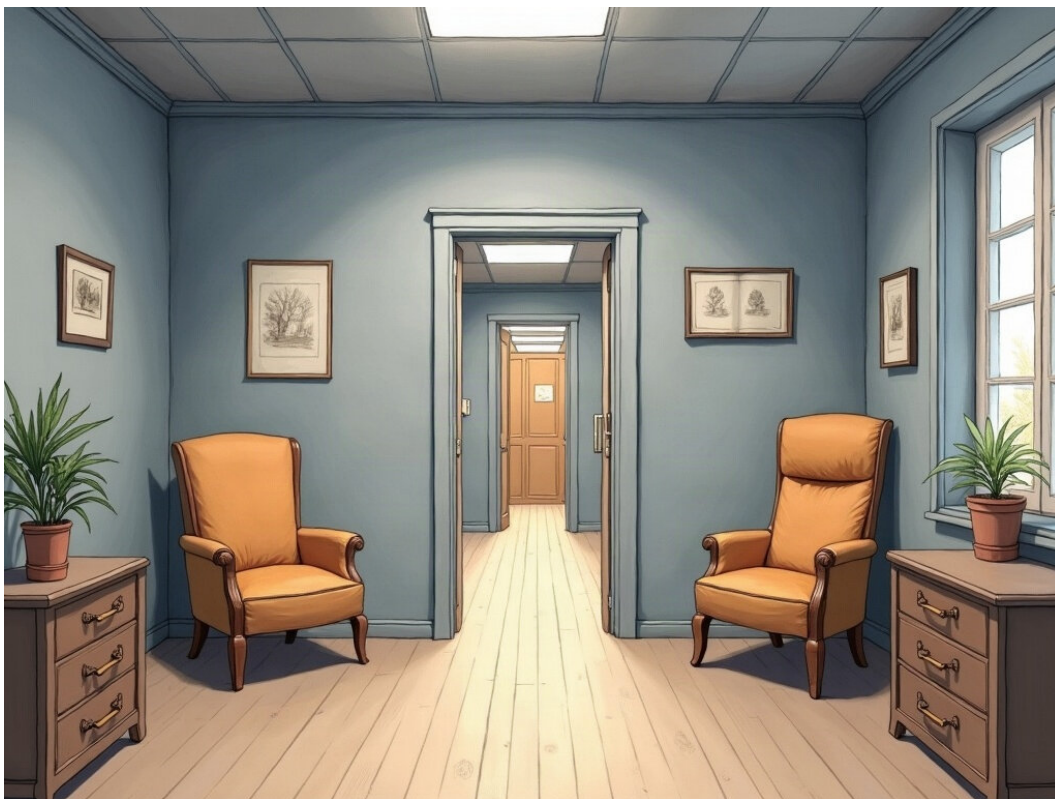
Так вопрос перестаёт быть игрой и становится практическим.

Квалиа

Слово, которым философы обозначают то, каково нечто переживать изнутри: краснота красного, вкус кофе, характер зубной боли. От латинского *qualis*, «какой». Главная их особенность — недоступность извне: сколько бы мы ни знали о длинах волн и работе глаза, из этого описания не следует, каково видеть красный. Существование квалиа — главное препятствие для объяснения сознания через одну только физику, и главная точка спора.

Комната Марии

Есть мысленный эксперимент, показывающий разрыв особенно наглядно. Придумал его австралийский философ Фрэнк Джексон в 1982 году.



Можно знать о цвете всё и ни разу его не увидеть

Мария — выдающийся учёный, изучающий зрение. Она знает **всё** о цвете: как устроен свет, как работают колбочки, какие нейроны возбуждаются, что происходит в мозге при виде красного. Абсолютно всё физически возможное знание.

При этом всю жизнь она прожила в чёрно-белой комнате и никогда не видела цвета.

Однажды её выпускают наружу, и она впервые видит красный помидор.

Вопрос: **узнала ли она что-то новое?**

Если да — значит, существует знание, не содержащееся в полном физическом описании. Значит, переживание не сводится к сведениям о нём.

Если нет — значит, полного физического описания достаточно, и «каково это» есть просто способ говорить об устройстве.

Большинство людей на первый вопрос отвечает «да, конечно узнала». И это отвечает многое: наша собственная интуиция настаивает, что переживание — нечто помимо описания.

Возражений к эксперименту много: одни говорят, что Мария не узнала новый факт, а приобрела новую **способность** — узнавать красное; другие — что мы просто не умеем вообразить по-настоящему полное знание. Спор идёт сорок лет.

Что с животными

Теперь туда, где вопрос перестаёт быть отвлечённым.

Собака скулит, когда ей больно. Рыба дёргается, когда её ранят. Насекомое убегает от опасности. Где кончается «реагирует» и начинается «чувствует»?

Вопрос не решается наблюдением, потому что наблюдаем мы только реакцию. Но косвенные признаки набрать можно, и биологи этим занимаются.

Схожее устройство. У млекопитающих те же структуры мозга и те же вещества, что участвуют в боли у нас. Довод сильный, но он работает хуже по мере удаления: у осьминога нервная система устроена совершенно иначе, и отсутствие наших структур ничего не доказывает.

Гибкое поведение, а не рефлекс. Существо, испытывающее боль, должно не просто отдёргиваться, а менять поведение надолго: беречь повреждённое место, избегать опасного места, выбирать обезболивание. Рыбы и раки в опытах демонстрируют именно такое.

Размен. Признак, который считают особенно весомым: животное **жертвует** чем-то ради облегчения. Краб покидает убежище, чтобы избежать удара током. Это больше, чем рефлекс: происходит взвешивание.



Где проходит граница между «реагирует» и «чувствует»

Заметь, как изменился разговор. Мы перестали искать доказательство и стали набирать вероятность. Это и есть разумный выход из положения, где доказательство невозможно.

И вывод отсюда практический: **осторожность разумна именно из-за неопределённости**. Не зная точно, чувствует ли существо, дешевле вести себя так, будто чувствует, — цена ошибки несимметрична.

Что с машинами

Теперь случай, который ещё десять лет назад был отвлечённым, а теперь нет.

Программа, с которой можно переписываться, говорит: «мне это неприятно». Она составляет фразы о своих состояниях так же гладко, как человек.

Довод по аналогии здесь ломается полностью. Программа ведёт себя похоже на человека, но устроена совершенно иначе: у неё нет тела, нет нервной системы, нет эволюционной истории, в которой боль была бы полезна. Сходство только в выходе, а не в устройстве.

Причём обрати внимание на неприятную деталь. Программа училась на текстах людей, описывающих свои чувства. То есть она умеет говорить о переживаниях **именно потому**, что копировала тех, у кого они есть. Её высказывание о боли — не свидетельство боли, а свидетельство того, что она хорошо усвоила, как о боли пишут.

Это не доказывает, что там ничего нет. Это показывает, что **поведение в данном случае вообще перестало быть уликой** — а другой улики у нас никогда и не было.

Можно ли вообще узнать, чувствует ли кто-то

Нельзя, и это принципиально. Наблюдению доступно только внешнее. Сколько бы мы ни изучали мозг, мы получим описание процессов, а не переживание. Между описанием и переживанием — разрыв, который не закрывается количеством сведений, и комната Марии это показывает.

Можно, просто пока не умеем. Разрыв кажется непреодолимым потому, что мы находимся в начале пути. Раньше казалось непостижимым, что такое жизнь, — и её объяснили химией без всякого особого вещества. Возможно, с сознанием будет так же: когда мы поймём, как именно оно возникает, вопрос «а вдруг там пусто» покажется таким же странным, как вопрос «а вдруг эта живая клетка не по-настоящему живая».

Спор этот — один из немногих, где обе стороны согласны, что решающего довода пока нет ни у кого.

Зачем вообще нужны переживания

Есть ещё один ход к той же задаче, и он подступает с неожиданной стороны: спросить не «есть ли они», а «зачем они».

Австралийский философ Дэвид Чалмерс в 1990-е предложил различать две задачи о сознании.

Лёгкие задачи — объяснить, как мозг различает цвета, управляет вниманием, хранит память, отчитывается о своих состояниях. Лёгкие

они не потому, что простые: это работа на столетия. Но понятно, **в каком виде** будет ответ: описание механизма.

Трудная задача — объяснить, почему всё это сопровождается переживанием. Почему обработка света не происходит в темноте, без всякого «каково это».

Ведь вообразить существо, которое делает ровно то же самое и при этом ничего не чувствует, как будто можно. Оно различало бы цвета, отдёргивало руку, говорило «мне больно» — и внутри у него было бы пусто. Философы называют такое существо зомби, и весь вопрос в том, возможно ли оно хотя бы мысленно.

Если возможно — значит, переживание не следует автоматически из устройства, и надо объяснять, откуда оно берётся.

Если невозможно — значит, достаточно сложное устройство неизбежно влечёт переживание, и трудной задачи попросту нет.

Обрати внимание, что мы вернулись к началу. Зомби — это в точности тот другой человек, о котором мы спрашивали в первом разделе, только рассмотренный не с вопросом «как узнать», а с вопросом «возможен ли он вообще». Первый вопрос про знание, второй про устройство мира, и второй, похоже, труднее.

Почему мы всё-таки не сомневаемся

Осталось объяснить, почему в обычной жизни вопрос не возникает.

Сомнение не работает практически. Невозможно всерьёз относиться к людям как к пустым оболочкам: ты не сможешь разговаривать, дружить, договариваться. Поведение, основанное на таком сомнении, разрушает всякое взаимодействие.

Отклик возникает раньше рассуждения. Увидев, как кто-то ударился, ты морщишься до всякого вывода. Это не умозаключение, а устройство: у людей есть встроенный механизм отклика на чужое состояние.

И это разумно. Вид, члены которого не считывали бы состояния друг друга, не выжил бы. Наша уверенность в чужих переживаниях — не философский вывод, а рабочее оснащение.

Философский вопрос при этом никуда не исчезает. Просто ответ на него даётся не доказательством, а тем, что иначе жить нельзя. Это слабое основание логически и очень прочное практически.

Что из этого следует

Несколько вещей, применимых вне философии.

Чужое состояние всегда реконструкция. Ты не видишь, что чувствует другой, — ты строишь предположение по внешним признакам. Значит, ошибаться будешь регулярно, и «я же вижу, что тебе всё равно» — почти всегда самонадеянность.

Слова — не доказательство, но лучшее, что есть. Поэтому спрашивать надёжнее, чем истолковывать. Не заменяет наблюдение, но резко уменьшает ошибку.

Неопределённость — основание для осторожности, а не для безразличия. Правило работает и с животными, и с людьми, чьё состояние тебе непонятно.

И самое трезвое. Единственное сознание, к которому у тебя есть прямой доступ, — твоё. Всё остальное принято на веру, основанную на сходстве. Помнить об этом полезно: это делает чуть скромнее в суждениях о том, что творится у других внутри.

Главное в трёх строчках

- Наблюдать можно только поведение: чужое переживание недоступно наблюдению в принципе, и вывод о нём строится по аналогии с единственным известным изнутри случаем — собой.
- Чем меньше сходства с нами, тем слабее этот вывод, поэтому вопрос из отвлечённого становится практическим применительно к животным и к машинам.
- Программа говорит о чувствах потому, что училась на текстах тех, у кого они есть, — значит, поведение перестало быть уликой, а другой улики у нас никогда и не было.

Вопросы для размышления

- Узнала ли Мария что-то новое, впервые увидев красное, — и что следует из твоего ответа?
- Почему довод «он устроен как я, значит чувствует как я» слабеет по мере удаления от человека, а не обрывается сразу?
- Если поведение перестало быть надёжной уликой, по какому признаку вообще стоило бы судить о наличии переживаний?