

**Keywords:** Information system, information garbage, quality of life, setization (networking), globalization.

#### References:

1. Boytsov B.V., Kryanev Yu.V. Implementation approach in the concept of quality of management. Theoretic – methodological aspects. *Quality and Life*. No. 4(12). 2016. Moscow, Pp. 7–13.
2. Yusupov R.M. Science and national security. St. Petersburg. Science. 2006. Available at: <http://www.myshared.ru/slide/337546/> (date of the address: 2/20/2017).
3. Yusupov R.M., Zabolotsky V.P. Scientific and methodological bases of informatization. Russian Academy of Sciences, Saint-Petersburg. *Institute of informatics and automation. Science*. 2000. 456 p.

4. Boytsov B.V., Kuznetsov M.A., Elkin G.I. Concept of quality of life: problems in global and regional measurements. *Academy of problems of quality*. 2007. Moscow, 238 p.
5. <https://tproger.ru/translations/sql-nosql-database-models/> (accessed: 22.05.2017).
6. <http://hr-portal.ru/blog/chto-takoe-bolshie-dannye-big-data> (accessed: 12.04.2017).
7. <http://lpgenerator.ru/blog/2015/11/17/chto-takoe-big-data-bolshie-dannye-v-marketinge-problemy-algoritmy-metody-analiza/> (accessed: 10.04.2017).
8. <https://cont.ws/@azovchanin77/582706> (accessed: 11.04.2017).
9. Borisenko A.A. Theory of systems. Information approach: monograph. Publishing house of Sumy state university. 2010. 210 p.

## Каузальные объяснения эффекта «зловещей долины» в робототехнике: теории и исследовательские данные

**И.В. Катерный**

к.филос.н., доцент МГИМО МИД России,  
старший научный сотрудник Института  
социологии ФНИСЦ РАН; Москва

e-mail: [yarkus@mail.ru](mailto:yarkus@mail.ru)

**Аннотация.** Когнитивное восприятие роботов и аватаров в условиях активного развития гибридных коммуникаций становится важной сферой исследования человека и общества – от нейронауки до философии. Приписывание «когнитивного» статуса (когнитивных состояний) неоморфным объектам в повседневном взаимодействии влияет на то, станут ли роботы полноценными социальными ко-агентами по примеру того, как это уже происходит с животными и человеческими эмбрионами. Основная проблема, которая стала всесторонне изучаться в когнитивных исследованиях, начиная с 2005 г., касается восприятия т.н. андроидных роботов, максимально приближенных к образу и поведению человека. Эта проблема, получившая известность как феномен «зловещей долины», открывает возможность проникнуть не только в эволюционные структуры человеческого сознания, но и понять механизмы воспроизводства социального порядка в целом. В статье предлагается обзор социальных, культурных и когнитивных теорий происхождения эффекта «зловещей долины» и обобщается десятилетний опыт научных исследований в этой области.

**Ключевые слова:** зловещая долина, зловещий обрыв, зловещая стена, андроидные роботы, гуманоидные роботы, теория управления ужасом, теория анимализма, теория негативных установок, теория эмпатии, теория нарушения ожиданий, морфотаксис.

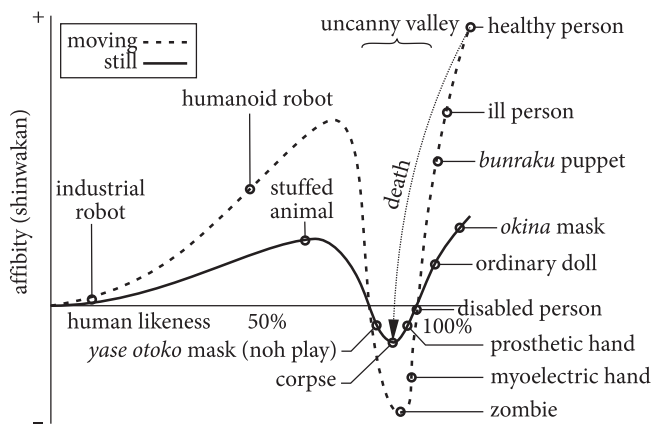
Когда в 1919 г. З. Фрейд в небольшом эссе [1] о немецких сказках под названием «*Unheimlich*» («жуткое», «зловещее») исследовал психоаналитические истоки чувства страха смерти («все то, что связано со смертью, с трупами и возвращением мертвых, с духами и призраками») и ужаса перед ее отдельными жуткими проявлениями («отрубленные члены, отсеченная голова, отрезанная от руки кисть, как в сказке Гауфа, ноги, пляшущие сами по себе, как в книге А. Шеффера, имеют в себе нечто необычайно зловещее, особенно когда они, как в последнем примере, наделяются еще и какой-то самостоятельной активностью»), он и не подозревал, как скоро все эти сказочные герои и события шагнут в реальную жизнь.

В 1970 г. японский робототехник Масахиро Мори на основе личного интроспективного наблюдения открыл и описал явление необычного психологического эффекта отторжения при восприятии искусственных объектов, максимально реплицирующих живого человека, а также его части тела. В частности, это касалось получивших в то время распространение довольно реалистичных миоэлектрических рук и протезов. Для небольшой заказной статьи о гуманоидной робототехнике, эра которой как раз начиналась в Японии, он построил гипотетический график, который, по его мнению, должен был отражать связь человекоподобности и перцептивной привлекательности разных субъектов и объектов, включая роботов. Гипотеза Мори состояла в утверждении, что чем сильнее объект похож на человека, тем более сильную симпатию он вызывает, но лишь до определенного предела, по-



сле которого возникает резкий спад. Этот эффект Мори, апеллируя к классической работе Фрейда, обозначил как «зловещая долина» (*bukimi no tani*). Более тридцати лет о работе Мори никто не слышал на Западе, и только с середины 2000-х гг., когда развитие робототехники по всему миру стало одним из главных направлений научно-технического творчества, возник настоящий бум интереса к этому феномену. Этому способствовал и первый перевод статьи Мори на английский язык, предпринятый американским исследователем из Университета Индианы К. МакДорманом [23]. Конструкторы и дизайнеры, занимающиеся созданием виртуальных персонажей и механических роботов «по образу и подобию» человека, узнали об интересном графике, созданном в свое время Мори (рис. 1).

Кардинальное отличие зоны «зловещей долины», изображенной на рис. 1 (провал), от крайней левой и крайней правой сторон графика заключается в том, что она включает в себя странные и гибридные создания, которые обладают смешанными



**Рис. 1. Оригинальный график М. Мори 1970 г.** (в переводе К. МакДормана) описывает связь между похожестью объекта на человека (ось абсцисс) и его привлекательностью (ось ординат). Чем более нечто похоже на живого человека, тем оно привлекательнее: движущийся гуманоидный робот (*humanoid robot*) и большая плюшевая игрушка в виде животного (*stuffed animal*) гораздо привлекательнее промышленного робота (*industrial robot*). Но на определенной фазе восприятия похожесть оборачивается отторжением и ужасом: маска демона (*yase otoko*) или старика (*okina*) из традиционного японского музыкального театра Но, реалистичная детская кукла (*ordinary doll*), человеческий труп (*corpse*), протез руки (*prosthetic hand*). При этом движущиеся объекты более подвержены эффекту «зловещей долины»: зомби (*zombie*), миоэлектрическая рука (*myoelectric hand*), марионетка (*bunraku puppet*) из традиционного японского кукольного театра Бунраку (способная моргать, высовывать язык, двигать бровями и пальцами). Пунктирная стрелка сверху вниз в нижнюю зону «зловещей долины» описывает ситуацию восприятия неожиданной смерти здорового человека [24]

характеристиками как «нормального» человека, так и неодушевленного (по некоторым своим проявлениям) объекта. Таким образом, данный график иллюстрирует предположение, что неожиданные резкие движения, неестественный цвет искусственной кожи, «мертвый» взгляд и другие детали поведения и внешности человекоподобного объекта способны вызвать страх и ужас, т.к. это *сигнализирует* о том, что перед нами не живое существо, а нечто чужеродное и опасное. Те же самые чувства могут возникать у человека при виде восковых кукол, зомби, покойников, одержимых, физических и умственных инвалидов и даже людей в припадке эпилепсии.

Конечно же, исследователи со всего мира принялись выяснять, существует ли валидный массив данных, способный подтвердить или опровергнуть ценность выдвинутой Мори гипотезы относительно восприятия неоморфов – гуманоидных (т.е. напоминающих человека) и особенно андроидных (т.е. максимально копирующих образ человека) роботов и аватаров. На момент окончания 2016 года были проведены десятки междисциплинарных эмпирических исследований, результаты которых весьма впечатляют, но единообразного научного ответа, тем не менее, не дают, т.к. разные гипотезы, методики и выборки в исследованиях приводили к различным выводам. Часть исследователей подтверждали существование нелинейного восприятия неоморфных объектов у взрослых людей [6, 10, 15, 21, 31, 33, 37], у детей [36], у младенцев [12, 18] и даже у приматов [32]. Другие исследования подтверждают альтернативные модели восприятия «жуткого», например, через паттерны «зловещего обрыва» (*uncanny cliff*) [2] или «зловещей стены» (*uncanny wall*) [34], а также отрицают усиление отталкивающего впечатления от восприятия движущихся андроидов [27, 38], что подразумевалось классической гипотезой Мори. В то же время, достаточных данных, полностью отрицающих наличие перцептивных отклонений в коммуникациях с андроидами, нет. Другими словами, вероятность возникновения «жуткого» впечатления от робота в той или иной степени сохраняется всегда. Более того, учет эффекта «зловещей долины» стал принципиально важен в области цифровой кинематографии, игровой индустрии и анимации после нескольких премьерных неудач в 2000-х гг. Стремление продюсеров максимально оживить цифровые образы людей (например, с помощью технологии «*motion-capture*») приводил к тому, что при восприятии их на экране у зрителя, особенно детского возраста, возникало неосознанное отторжение. Мелкие аномалии, которые сопровождают мимику, жестикуляцию и эмоциональное поведение

ние слишком похожего на обычного человека персонажа (например, т.н. эффект «мертвого глаза»), делают его «жутким». В качестве примеров исследователи приводят неудачные образы, созданные в «Последней фантазии» (2001), «Полярном экспрессе» (2004), «Беовульфе» (2007), «Приключениях Тин-Тина» (2011) [35].

**Каковы же основные возможные причины вероятного проявления эффекта «зловещей долины» при контакте с андроидами?**

*Теория управления ужасом.* Неспроста одной из первых гипотез, претендовавших на объяснение феномена «зловещей долины», стала теория управления страхом (ужасом), разработанная в 1980-х гг. социальными психологами Дж. Гринбергом, Ш. Солломоном и Т. Пышчински [8] на основе идей известного американского культурного антрополога и писателя Э. Беккера [4]. В ее основе лежит гипотеза «мортальной тревожности» (*mortal salience hypothesis*), которая подразумевает кардинальное знание человека о неизбежности своей смерти, но в то же время предполагает его кардинальное желание избегать мыслей и напоминаний об этом: сознательные мысли о смерти рационализируются (здоровый образ жизни и т.п.) или подавляются, неосознаваемые сигналы о смерти приводят в действие защитные механизмы, направленные на поддержание значения интегрирующих и жизнеутверждающих элементов культуры и личности (самоуважения) в обществе. В этом смысле страх смерти может быть или проксимальным (от осознаваемых прямых угроз), или дистальным (от непрямых неосознаваемых угроз), например со стороны враждебной пропаганды. Теория управления страхом на основе сотен проведенных экспериментов утверждает, что если базовая инкультурация позволяет действительно избавиться от постоянной мортальной тревожности (через прямые и символические объяснения возможности бессмертия), то дистальные угрозы через защитные психические механизмы будут вызывать усиление приверженности усвоенным культурным ценностям, что в частности, проявляется в предпочтении верить больше положительным, а не отрицательным оценкам своей страны [9], склоняться к выбору харизматических, а не демократических лидеров [7] и строго осуждать нарушения моральных норм [29]. Восприятие (на изображении) андроидных роботов также может иметь отношение к дистальным угрозам, поэтому одним из первых экспериментов, направленных на эмпирическое выявление эффекта «зловещей долины», был эксперимент, направленный на выявление связи между «жуткостью» изображений андроидов и проявлениями подсознательного управления

дистальным страхом смерти [13]. В итоге, экспериментальная группа, которая была подвергнута испытанию вызывания сублиминальной мортальной тревожности (изображением на экране компьютера головы и торса андроидного робота), действительно показала, отвечая после этого на вопросы специальной анкеты, более высокие показатели предпочтений харизматически окрашенных политических речей и привлекательности благоприятно отзывающегося об их стране иностранце (по отношению к более критически настроенному), по сравнению с контрольной группой, которая не видела предположительно «жуткие» изображения и тем самым не подвергалась мортальной тревоге. Однако, как указали сами экспериментаторы, результаты вышли все-таки не очевидными, т.к. оцениваемые эффекты дистального страха не могут быть надежным индикатором мортальной тревоги для изучения эффекта «зловещей долины» в виду его перцептуальной специфичности. Восприятие таких образов слишком сильно зависит от длительности, ситуации и индивидуальных аффективных реакций, которые не принимаются в расчет теорией управления ужасом. Таким образом, самый главный вопрос о том, что конкретно вызывает чувство «жуткости» в андроидах, остался без ответа и открывал простор для новых гипотез [*Ibid.*, с. 115].

*Теория нарушения ожиданий.* К. МакДорман и его коллега из Японии известный инженер Х. Исигуро для объяснения эффекта «зловещей долины» предложили гипотезу, которая развивает понятие «ошибки предсказания» и предполагает, что коммуникация с андроидами нарушает паттерн микроповеденческих взаимодействий, где любые действия становятся результатом реализации правильно понятых ожиданий, которые в свою очередь производятся и приписываются партнеру. Андройды нарушают эту связь из-за перцептуальных и сенсомоторных недостатков в своем поведении, и нормальная модель человеческой интеракции разрушается. Другими словами, чем более критичны для человека ошибки в поведении робота, тем более он склонен к восприятию «жуткого» в нем [14].

*Теория эмпатии.* Эффект «зловещей долины» может возникать в ситуации, когда человек испытывает внутреннее напряжение (дистресс) в связи с такими проявлениями эмпатии, как дискомфорт или тревога. Андройды могут вызвать дистресс несколькими способами. Во-первых, поскольку внешне они выглядят как люди, то их специфические движения могут напоминать приступы эпилепсии или какие-нибудь физические нарушения, впечатление от которых ведет к внутреннему дискомфорту. Во-вторых, роботы могут вызвать тревожность из-за подсознательного страха перед самостоятель-





но двигающимися, но мертвыми созданиями, что, как указывал еще Фрейд, является одним из наиболее сильных страхов человека. В-третьих, симпатия по отношению к роботу подсознательно блокируется из-за осознания того факта, что машина – это не человек и не живое создание, т.е. не отвечает концепции моральной феноменологии [19].

*Теория «анимализма».* Группа нейрофизиологов из Калифорнийского института технологий в 2011 году смогла обнаружить область в человеческом мозге, отвечающую исключительно за восприятие животных [25]. Этот своего рода «звериный плагин», эволюционно возникший на заре человечества, отвечает за такие эмоции как страх и ужас перед угрожающими жизни животными (медведи и т.п.). До этого было установлено, что еще одной эволюционной эмоцией, позволяющей человеку выживать, является отвращение к потенциально патогенным организмам, несущим угрозу заражения. Все отвратительные (для человека) объекты имеют животное (природное) происхождение (испражнения, тараканы, разлагающиеся тела и т.п.) [30, с. 22]. И в том, и другом случае действуют подсознательные сигналы об опасности, что также может быть применимо к восприятию неоморфных объектов, которые из-за своей странности могут выглядеть и вести себя опасно и отвратительно. Чувствительность к стимулированию анималистических угроз при виде андроидов также может создавать эффект «зловещей долины», напоминая человеку о его собственной биологичности, животности и смертности, что в свою очередь уже приводит в действие механизм управления ужасом.

*Теория когнитивного диссонанса.* На заре эмпирических исследований феномена «зловещей долины» в 2005 г. американский исследователь К. Рэйми высказал идею, что восприятие андроидов сталкивается с проблемой категоризации: две качественно различные категории – человек и робот – оказываются смешанными в количественных параметрах внешности и поведения (уровне человекоподобия) андроида [28]. Похожую картину, по Рэйми, можно наблюдать в дискуссиях о запрете аборт, где вопрос о том, с какого времени считать эмбрион живым человеком, остается неурегулированным. Когнитивная неопределенность в отношении андроидов может быть снята либо введением третьей категории, в которую будут входить все двойственные неоморфные создания, либо снятием базовой дихотомии и усвоением концепции онтологического единства всех наличных созданий как сложных физических механизмов. Однако перцептуальная (а не сугубо рациональная) природа «жуткого» делает эту задачу затруднительной и ставит вопрос о происхождении базовой бинарности «человек – машина».

*Теория негативных установок.* Популяризация достижений робототехнической отрасли имеет в качестве оборотной стороны распространение мнения об угрожающем характере роботизации жизни, что охотно педалируется образами массовой культуры («терминаторы» и т.п.). Негативное, эмоционально окрашенное отношение к андроидам может быть вызвано именно мировоззренческими принципами наподобие экологических установок. Даже в Японии, где, как считается, все любят роботов, на самом деле существует проблема эмоционального сближения с человекоподобными искусственными созданиями и люди скорее предпочитают иметь дело с людьми [3, с. 16]. Существует специальная анкета, измеряющая (по шкале Лайкерта) степень негативного отношения к роботам по трем основаниям: отношение к ситуациям взаимодействия с роботами (например, согласие/несогласие с утверждением «я бы чувствовал себя неловко на работе, где необходимо использовать роботов»); отношение к воздействию роботов на общество (например, согласие/несогласие с утверждением «я боюсь, что роботы будут доминировать в обществе будущего») и отношение к эмоциональному контакту с роботом (например, согласие/несогласие с утверждением «если бы у робота были бы эмоции, я бы запросто подружился с ним») [26]. Негативная эмоциональная составляющая установки на взаимодействие с роботами как раз может вызывать страх, тревогу и быть причиной проявления эффекта «зловещей долины».

*Теория религиозного фундаментализма.* Данная гипотеза развивает предыдущую теорию, но в более узком отношении. Предполагается, что ортодоксально настроенные представители авраамических религий более склонны, чем другие люди (атеисты или синтоисты), видеть андроидов жуткими [17]. Это может быть связано с мировоззрением, которое настроено негативно к техническому и социальному прогрессу и догматически резко различает онтологическую область существования человеческих существ и всю остальную сферу бытия, включая животных и искусственные создания. Андроиды, как негативная референтная группа (наравне с гомосексуалистами, язычниками, расистами и т.п.), угрожает онтологической безопасности человека. Помимо этого существование андроидов подсознательно разрушает концепцию «души» как источника жизни: машина, принимающая облик и роль человека, делает понятие человеческой души избыточным. Более того, возможность замены и перестановки частей тела андроида подрывает концепцию эксклюзивного человеческого бессмертия, что способно вызывать страх, тревогу и ужас.

В одном из наиболее масштабных исследований пионера американской андроид-науки К. МакДормана (совместно со С. Энтезари) некоторые из этих теорий прошли операционализацию и были изучены на выборке около 600 (на разных этапах) человек на предмет их перцептуальной чувствительности [*Ibid.*]. Все участники сначала прошли тесты на подверженность тем или иным факторам, описанным в различных гипотезах (в основном по шкале Лайкерта), после чего просматривали видеоролики с шестью движущимися изображениями робота-пылесоса, гуманоидного робота, трех андроидов с бросающимися в глаза девиациями во внешности и человека. Эффект «зловещей долины» оценивался через индикаторы «жуткости», т.е. отталкивающего впечатления, и «теплоты».

Как видно из *табл. 1*, были проанализированы девять характеристик личности, интерпретированных как факторы чувствительности к эффекту «зловещей долины»: негативные установки по отношению к роботу; подсознательное отвращение («анимализм») в восприятии; диссонанс при категоризации андроида в когнитивной бинарности «человек – машина»; тревожность; невротичность; озабоченность ошибками в поведении («перфекционизм»); эмпатический дистресс; религиозный фундаментализм; когнитивная вы-

деленность андроидов в отдельную категорию окружающих объектов («человек – робот – андроид»). Поскольку принимались в расчет два индикатора, описывающие «зловещую долину» («жуткость» и «теплота»), то наиболее слабые корреляции показали факторы озабоченности ошибками и когнитивной категоризации «человек – робот – андроид».

После проведения межфакторного анализа были выявлены пять основных характеристик, которые имеют максимальные значения по индикаторам «зловещей долины» и высокие коэффициенты взаимной корреляции. Их причинно-следственная связь была смоделирована методом структурных уравнений (*рис. 2*).

Религиозный фундаментализм, «анимализм» и повышенная тревожность, чувствительные к девиациям во внешнем облике и поведении робота, оказывают косвенное влияние на эффект «зловещей долины», тогда как когнитивная приверженность жесткой категоризации «человек-машина» и негативные эмоции к роботам напрямую увеличивают жуткость или снижают теплоту в восприятии андроида. При этом разные факторы имеют неодинаково выраженное направление. Например, религиозный фундаментализм и перцептивный диссонанс сильнее коррелируют с индикатором (снижения) «теплоты», а «анимализм» и тревожность больше

Таблица 1

**Коэффициенты корреляции Пирсона между анализируемыми факторами «зловещей долины» и оценками «жуткости» и «теплоты» при восприятии андроидов; Альфа Кронбаха надежности анализируемого фактора с учетом количества полученных ответов; среднее и стандартное отклонение полученных значений [17]**

Факторы/Индикаторы	«Жуткость»	«Теплота»	Альфа Кронбаха	Среднее	Стандартное отклонение
Негативные установки по отношению к робототехнике	.41***	-.38***	.86 (n=575)	-0.11	1.65
Подсознательное отвращение к роботам	.34***	-.21***	.83(n=571)	-0.22	1.38
Перцептивный диссонанс при восприятии робота (смешение «живого» и «неживого»)	.28***	-.37***	.89 (n=573)	0.93	1.31
Тревожность из-за «необычного» поведения робота	.18***	-.13***	.87 (n=563)	-0.20	1.10
Страх и другие неврозы при контакте с роботами	.16***	-.12***	.77 (n=571)	-0.16	1.18
Перфекционизм: чувствительность при восприятии ошибок в поведении робота	.13**	.01	.88 (n=576)	-0.38	1.30
Эмпатический дистресс при восприятии «неполноценности» робота	.12**	-.07	.83 (n=577)	-1.01	1.10
Религиозный фундаментализм: неприятие роботов как «божьих созданий»	.12**	-.24***	.96 (n=563)	-1.51	2.26
Отношение к андроидам как особому виду роботов	-0.03	.26***	.75 (n=570)	0.53	1.04

\* p < .05. \*\* p < .01. \*\*\* p < .001.



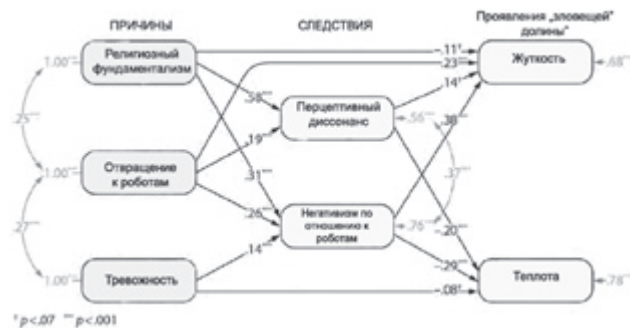
воздействуют именно на усиление «жуткого». По словам К. МакДормана, это можно связать с природой этих характеристик – первые из них являются культурными паттернами, а вторые представляют собой механизмы биологической адаптации к обнаружению внешней угрозы [Ibid., с. 160].

Хотя это исследование К. МакДормана и С. Энтезари встретило ряд критических оценок методологического характера [5], для нас более важным являются выводы, которые мы можем сделать в целом, исходя из опыта десятилетнего изучения феномена «зловещей долины».

Прежде всего, необходимо отметить, что исследование механизмов восприятия неоморфных объектов имеет куда большее значение для познания человека, нежели научное обеспечение дизайнерской работы по созданию привлекательных моделей новых роботов. Чем глубже мы погружаемся в тему, тем явственнее становится понимание того, что отторжение жуткого в человеческом восприятии является лишь поверхностным эффектом воспроизводства *морфотаксиса* – фундаментального порядка, который включает в себя и связывает между собой различные биологические, когнитивные и социальные и другие структуры. Прозопопический морфогенезис, связанный с активным вовлечением роботов и аватаров в социальное сообщество, сталкивается с сопротивлением ригидных морфотаксических элементов, встроенных в «базовую реальность» человека с целью обеспечения его выживания на физическом уровне и соответствующей институционализации норматив-

ных суперструктур. Особенностью морфотаксиса в этом смысле является его бинарная организация вокруг базовых эволюционных дихотомий, формирующих онтологические границы человеческого мира. Как показывают исследования, одними из таких онтологически кардинальных дихотомий являются (пересекающиеся) различия «человек – не-человек» и «живой – мертвый». Проблемы аборт, эвтанази, искусственного интеллекта и андроидных роботов, по сути, манифестируют онтологическую напряженность именно этих полюсов.

Эффект «зловещей долины» указывает на то, что в основе солидарности социального сообщества (как людей, так и приматов) лежат не только механизмы морального и символического признания, но и более глубокие, а точнее – более фундаментальные механизмы *непризнания*. «Жуткое» как феномен непризнания пронизывает морфотаксис на биологическом уровне (характерная нейронная активность), бессознательном (страх перед смертью), когнитивном (чувства отвращения и ужаса), социальном (стигматизация) и культурном (непризнание морального статуса). И как, в частности, выяснилось в ходе последнего описанного исследования, несмотря на то, что большая часть испытуемых были атеистами, в среднем все они поддержали тезис религиозного (а именно «авраамического») фундаментализма о том, что «люди фундаментально отличаются от роботов» [17, с. 161]. Таким образом, создание андроидов по ту сторону «зловещей долины» оказывается вопросом не просто техническим. В близком или отдаленном будущем технически это, наверняка, станет достижимым, но, как мудро заметил сам М. Мори в одном из интервью, отвечая на вопрос о том, возможно ли это в принципе, он ответил: «Да, но зачем это нужно?», – подчеркивая тем самым фундаментальную необходимость сохранения, а не стирания онтологического разрыва между роботом и человеком во имя его собственного выживания и поддержания социального и морального порядка [11].



**Рис. 2. Диаграмма путей в структурном уравнении причинно-следственных связей факторов, чувствительных к эффекту «зловещей долины» (слева направо): религиозный фундаментализм, отвращение к роботам («анимализм»), тревожность; перцептивный диссонанс, негативные эмоции по отношению к роботам. Индикаторы «зловещей долины»: (эмоциональная) жуткость, (эмоциональная) теплота. Стрелки указывают направление прямых причинно-следственных связей, дуги – непрямые связи. Цифры показывают относительные весовые значения оценки прямых и непрямые связей; звездочки – степень достоверности весовых значений связей [17]**

## Литература

1. Фрейд, З. Жуткое / З.Фрейд // Фрейд З. Собр. соч. Т.4: Психологические сочинения. – М.: «Фирма СТД», 2006. – С. 261–297.
2. Bartneck C., Kanda T., Ishiguro H., Hagita N. Is the Uncanny Valley an Uncanny Cliff? // Proceedings of the 16th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN. 2007, 26-29 August, Jeju, Korea. New York: IEEE, 2007. P. 368–373.



3. Bartneck C., Suzuki T., Kanda T., Nomura T. The Influence of People's Culture and Prior Experiences with Aibo on their Attitude Towards Robots // *AI and Society*. 2007. Vol.21. № 1/2. P. 217–230.
4. Becker E. *The Denial of Death*. New York: Simon and Schuster, 1973.
5. Burleigh T.J. A Challenge to the Study of Individual Differences in Uncanny Valley Sensitivity: The Importance of Looking at Individual-Level Response Patterns // *Interaction Studies*. 2015. Vol.16. № 2. P. 186–192.
6. Burleigh T.J., Schoenherr J.R., Lacroix G.L. Does the Uncanny Valley Exist? An Empirical Test of the Relationship between Eeriness and the Human Likeness of Digitally Created Faces. *Computers in Human Behavior*. 2013. Vol. 29. № 3. P. 759–771.
7. Cohen F., Solomon S., Maxfield M., Pyszczynski T., Greenberg J. Fatal Attraction: The Effects of Mortality Salience on Evaluations of Charismatic, Task-Oriented, and Relationship-Oriented Leaders // *Psychological Science*. 2004. Vol. 15. P. 846–851.
8. Greenberg J., Pyszczynski T., Solomon S. The Causes and Consequences of the Need for Self-Esteem: A Terror Management Theory // Baumeister R.F. (ed) *Public Self and Private Self*. New York: Springer-Verlag, 1986. P. 189–212.
9. Greenberg J., Pyszczynski T., Solomon S., Rosenblatt A., Veeder M., Kirkland S., Lyon D. Evidence for Terror Management Theory II: The Effects of Mortality Salience Reactions to Those who Threaten or Bolster the Cultural Worldview // *Journal of Personality and Social Psychology*. 1990. Vol. 58. P. 308–318.
10. Ho C.-C., MacDorman K.F., Dwi Pramono Z.A.D. Human Emotion and the Uncanny Valley: A GLM, MDS, and Isomap Analysis of Robot Video Ratings // *Proceedings of the 3d ACM/IEEE International Conference on Human–Robot Interaction*. 2008, March 12–15, Amsterdam, Netherlands. New York: ACM, 2008. P. 169–176.
11. Katayama L. How Robotics Master Masahiro Mori Dreamed Up the Uncanny Valley // *Wired*. 2011, 29 November. URL: [http://www.wired.com/2011/11/pl\\_mori/](http://www.wired.com/2011/11/pl_mori/) (дата обращения: 02.08.2016).
12. Lewkowicz D.J., Ghazanfar A.A. The Development of the Uncanny Valley in Infants // *Developmental Psychobiology*. 2012. Vol. 54. № 2. P. 124–132.
13. MacDorman K. Androids as an Experimental Apparatus: Why is There an Uncanny and Can We Exploit It? // *Proceedings of the CogSci-2005, XXVII Annual Conference of the Cognitive Science Society*. 2005, July 25–26, Stresa. Mahwah, NJ: Lawrence Erlbaum Ass., 2005. P. 108–118.
14. MacDorman K., Ishiguro H. The Uncanny Advantage of Using Androids in Social and Cognitive Science Research // *Interaction Studies*. 2006. Vol. 7. № 3. P. 297–337.
15. MacDorman K. F., Green R. D., Ho C.-C., & Koch C. Too Real for Comfort: Uncanny Responses to Computer Generated Faces // *Computers in Human Behavior*. 2009. Vol. 25. № 3. P. 695–710.
16. MacDorman K.F., Vasudevan S.K., Ho C.-C. Does Japan Really Have Robot Mania? Comparing Attitudes by Implicit and Explicit Measures // *AI and Society*. 2009. Vol. 23. № 4. P. 485–510.
17. MacDorman K.F., Entezari S.O. Individual Differences Predict Sensitivity to the Uncanny Valley // *Interaction Studies*. 2015. Vol.16. № 2. P.141–172.
18. Matsuda Y.T., Okamoto Y., Ida M., Okanoya K., Myowa-Yamakoshi M. Infants Prefer the Faces of Strangers or Mothers to Morphed Faces: The Uncanny Valley between Social Novelty and Familiarity // *Biology Letters*. 2012. Vol. 8. № 5. P. 725–728.
19. Misselhorn C. Empathy with Inanimate Objects and the Uncanny Valley // *Minds and Machines*. 2009. Vol.19. №.1. P. 345–359.
20. Mitchell W.J., Szerszen Sr K.A., Lu A.S., Schermerhorn P.W., Scheutz M., MacDorman K. F. A Mismatch in the Human Realism of Face and Voice Produces an Uncanny Valley // *i-Perception*. 2011. Vol. 2. № 1. P. 10–12.
21. Moosa M.M., Ud-Dean, S.M.M. Danger Avoidance: An Evolutionary Explanation of the Uncanny Valley // *Biological Theory*. 2010. Vol. 5 №1. P. 12–14.
22. Mori M. Bukimi No Tani [The Uncanny Valley] / trans. by K.F. MacDorman, T. Minato // MacDorman K. *Androids as an Experimental Apparatus: Why is There an Uncanny and Can We Exploit It?* (Appendix B) // *Proceedings of the CogSci-2005, XXVII Annual Conference of the Cognitive Science Society*. 2005, July 25–26, Stresa, Italy. Mahwah, NJ: Lawrence Erlbaum Ass., 2005. P. 108–118.
23. Mori M. The Uncanny Valley / trans. by K.F. MacDorman and N. Kageki // *IEEE Robotics and Automation*. 2012. Vol. 19. № 2. P. 98–100.
24. Mormann F., Dubois J., Kornblith S., Milosavljevic M., Cerf M., Ison M., Tsuchiya N., Kraskov A., Quiroga R., Adolphs R., Fried I., Koch Ch. A Category-Specific Response to Animals in the Right Human Amygdale // *Nature Neuroscience*. 2011. Vol. 14. № 10. P. 1247–1249.
25. Nomura T., Kanda T., Suzuki T. Experimental Investigation into Influence of Negative Attitudes toward Robots on Human–Robot Interaction // *AI and Society*. 2006. Vol. 20. № 2. P. 138–150.
26. Piwek L., McKay L.S., Pollick F.E. Empirical Evaluation of the Uncanny valley Hypothesis Fails to Confirm the Predicted Effect of Motion // *Cognition*. 2014. Vol. 130. № 3. P. 271–277.
27. Ramey C.H. The Uncanny Valley of Similarities Concerning Abortion, Baldness, Heaps of Sand, and Humanlike Robots // *Proceedings of the Workshop 'Views of the Uncanny Valley': 5th IEEE-RAS Interna-*



tional Conference on Humanoid Robots. 2005, December 05-07, Tsukuba, Japan. 2005. Los Alamitos: IEEE Computer Society Press. P. 8–13.

29. Rosenblatt A., Greenberg J., Solomon S., Pyszczynski T., Lyon D. Evidence for Terror Management Theory I: The Effects of Mortality Salience on Reactions to Those who Violate or Uphold Cultural Values // *Journal of Personality and Social Psychology*. 1989. Vol. 57. P. 681–690.

30. Rozin P., Fallon A.E. A Perspective on Disgust // *Psychological Review*. 1987. Vol. 94. P. 23–41.

31. Seyama J., Nagayama R.S. The Uncanny Valley: The Effect of Realism on the Impression of Artificial Human Faces // *Presence: Teleoperators and Virtual Environments*. 2007. Vol. 16. № 4. P. 337–351.

32. Steckenfinger S. A., Ghazanfar A. A. Monkey Visual Behavior Falls into the Uncanny Valley // *Proceedings of the National Academy of Sciences*. 2009. Vol. 106. № 43. P.18362–18366.

33. Tinwell A., Grimshaw M., Abdel Nabi D., Williams A. Facial Expression of Emotion and Per-

ception of the Uncanny Valley in Virtual Characters // *Computers in Human Behavior*. 2011. Vol. 27. № 2. P. 741–749.

34. Tinwell A., Grimshaw M., Williams A. The Uncanny Wall // *International Journal of Arts and Technology*. 2011. Vol. 4. № 3. P. 326–341.

35. Tinwell A. *The Uncanny Valley in Games and Animation*. Boca Raton, FL: CRC Press, 2014.

36. Tinwell A., Sloan R.J.S. Children's Perception of Uncanny Human-Like Virtual Characters // *Computers in Human Behavior*. 2014. Vol. 36. P. 286–296.

37. Tinwell A., Grimshaw M., Abdel Nabi D. The Effect of Onset Asynchrony in Audio Visual Speech and the Uncanny Valley in Virtual Characters // *International Journal of Mechanisms and Robotic Systems*. 2015. Vol. 2. № 2. P. 97–110.

38. Thompson J.C., Trafton J.G., McKnight P. The Perception of Humanness from the Movements of Synthetic Agents // *Perception*. 2011. Vol.40. № 6. P. 695–704.

### Causal Explanations of the Uncanny Valley Effect in Robotics: Theories and Applied Research

**I.V. Katernyi**, *Candidate of philosophical sciences, Associate Professor at MGIMO-University, Senior researcher at Institute of Sociology of the Federal Center of Theoretical and Applied Sociology of the Russian Academy of Sciences; Moscow*

*e-mail: yarkus@mail.ru*

**Summary.** Cognitive perception of robots and avatars in the conditions of active development of hybrid communications becomes the important sphere of a research of the person and society – from neuroscience to philosophy. Attributing of the «cognitive» status (cognitive states) to neomorphic objects in daily interaction influences whether robots will be full social co-agents alongside animals, pets and human embryos. The main problem which began to be studied comprehensively in cognitive studies since 2005, concerns deviations when communicating with android robots constructed to have humanlike appearance and behavior as much as possible. This problem which has become popularized as the «uncanny valley» effect opens new avenues to get not only into evolutionary structures of human consciousness, but also to understand mechanisms of reproduction of a social order in general. In the article, some relevant theories to explain the uncanny phenomenon and a related body of research data are provided.

**Keywords:** uncanny valley, uncanny cliff, uncanny wall, android robots, humanoid robots, terror management theory, animal reminder sensitivity, negative attitudes to robots.

### References:

1. Freud Z. Terrible. Collected works. V. 4: Psychological compositions. *Firm STD*. 2006. Moscow, pp. 261–297.

2. Bartneck C., Kanda T., Ishiguro H., Hagita N. Is the Uncanny Valley an Uncanny Cliff? Proceedings of the 16th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN. 2007. Korea. *New York: IEEE*. 2007. pp. 368–373.

3. Bartneck C., Suzuki T., Kanda T., Nomura T. The Influence of people's culture and prior experiences with aibo on their attitude towards robots. *AI and Society*. 2007. V. 21. pp. 217–230.

4. Becker E. *The denial of death*. *New York: Simon and Schuster*. 1973.

5. Burleigh T.J. A Challenge to the study of individual differences in uncanny valley sensitivity: The importance of looking at individual-level response patterns. *Interaction Studies*. 2015. V. 16. No. 2. P. 186–192.

6. Burleigh T.J., Schoenherr J.R., Lacroix G.L. Does the uncanny valley exist? An empirical test of the relationship between eeriness and the human likeness of digitally created faces. *Computers in Human Behavior*. 2013. V. 29. No. 3. pp.759–771.

7. Cohen F., Solomon S., Maxfield M., Pyszczynski T., Greenberg J. Fatal attraction: the effects of mortality salience on evaluations of charismatic, task-oriented, and relationship-oriented leaders. *Psychological science*. 2004. V. 15. pp. 846–851.

8. Greenberg J., Pyszczynski T., Solomon S., Baumeister R.F. The causes and consequences of the need for self-esteem: A terror management theory. *Public Self and Private Self*. 1986. New York: Springer-Verlag, pp. 189–212.

9. Greenberg J., Pyszczynski T., Solomon S., Rosenblatt A., Veeder M., Kirkland S., Lyon D. Evidence for terror management theory II: The effects of mortality



salience reactions to those who threaten or bolster the cultural worldview. *Journal of personality and social psychology*. 1990. V. 58. pp. 308–318.

10. Ho C.-C., MacDorman K. F., Dwi Pramono Z.A.D. Human Emotion and the Uncanny Valley: A GLM, MDS, and isomap analysis of robot video ratings. Proceedings of the 3d ACM/IEEE international conference on human-robot interaction. *New York: ACM*. 2008. Amsterdam, Netherlands, pp. 169–176.

11. Katayama L. How robotics master Masahiro Mori dreamed up the uncanny valley. *Wired*. 2011, Available at: [http://www.wired.com/2011/11/pl\\_mori/](http://www.wired.com/2011/11/pl_mori/) (accessed: 02.08.2016).

12. Lewkowicz D.J., Ghazanfar A. A. The Development of the uncanny valley in infants. *Developmental Psychobiology*. 2012. V. 54. No. 2. pp. 124–132.

13. MacDorman K. Androids as an experimental apparatus: Why is there an uncanny and can we exploit it? Proceedings of the cogSci-2005, XXVII annual conference of the cognitive science society. *Stresa, Mahwah, NJ: Lawrence Erlbaum Ass.* 2005. pp. 108–118.

14. MacDorman K., Ishiguro H. The Uncanny advantage of using androids in social and cognitive science research. *Interaction studies*. 2006. V. 7. No. 3. pp. 297–337.

15. MacDorman K. F., Green R. D., Ho C.-C., & Koch C. Too real for comfort? uncanny responses to computer generated faces. *Computers in human behavior*. 2009. V. 25. No. 3. pp. 695–710.

16. MacDorman K.F., Vasudevan S.K., Ho C.-C. Does Japan really have robot mania? Comparing attitudes by implicit and explicit measures. *AI and society*. 2009. V. 23. No. 4. pp. 485–510.

17. MacDorman K.F., Entezari S.O. Individual differences predict sensitivity to the uncanny valley. *Interaction Studies*. 2015. V. 16. No. 2. pp.141–172.

18. Matsuda Y. T., Okamoto Y., Ida M., Okanoya K., Myowa-Yamakoshi M. Infants prefer the faces of strangers or mothers to morphed faces: The uncanny valley between social novelty and familiarity. *Biology letters*. 2012. V. 8. No. 5. pp. 725–728.

19. Misselhorn C. Empathy with inanimate objects and the uncanny valley. *Minds and machines*. 2009. V. 19 №.1. pp. 345–359.

21. Mitchell W.J., Szerszen Sr K.A., Lu A.S., Schermerhorn P.W., Scheutz M., MacDorman K.F. A mismatch in the human realism of face and voice produces an uncanny . *I-Perception*. 2011. V. 2. No. 1. pp. 10–12.

22. Moosa M.M., Ud-Dean, S.M.M. Danger avoidance: an Evolutionary explanation of the uncanny valley. *Biological theory*. 2010. V. 5 No. 1. pp. 12–14.

23. Mori M. Bukimi No Tani, MacDorman K.F., Minato T. [The Uncanny Valley] Androids as an experimental apparatus: Why is there an uncanny and can we exploit It? Proceedings of the CogSci-2005, XXVII Annual conference of the cognitive science society. Italy. Mahwah, NJ: Lawrence Erlbaum Ass., 2005. pp. 108–118.

24. Mori M., Kageki N., MacDorman K.F., The uncanny valley. *IEEE Robotics and automation*. 2012. V. 19. No. 2. pp. 98–100.

25. Mormann F., Dubois J., Kornblith S., Milosavljevic M., Cerf M., Ison M., Tsuchiya N., Kraskov A., Quian Quiroga R., Adolphs R., Fried I., Koch Ch. A category-specific response to animals in the right human amygdale. *Nature neuroscience*. 2011. V. 14. No. 10. pp. 1247–1249.

26. Nomura T., Kanda T., Suzuki T. Experimental investigation into influence of negative attitudes toward robots on human-robot interaction. *AI and society*. 2006. V. 20. No. 2. pp. 138–150.

27. Piwek L., McKay L.S., Pollick F.E. Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion. *Cognition*. 2014. V. 130. No. 3. pp. 271–277.

28. Ramey C.H. The uncanny valley of similarities concerning abortion, baldness, heaps of sand, and humanlike robots. Proceedings of the Workshop 'Views of the uncanny valley': 5th IEEE-RAS International conference on humanoid robots. *IEEE Computer Society Press*. 2005. pp. 8–13.

29. Rosenblatt A., Greenberg J., Solomon S., Pyszczynski T., Lyon D. Evidence for terror management theory I: The effects of mortality salience on reactions to those who violate or uphold cultural values. *Journal of personality and social psychology*. 1989. V. 57. pp. 681–690.

30. Rozin P., Fallon A.E. A Perspective on disgust. *Psychological review*. 1987. V. 94. pp. 23–41.

31. Seyama J., Nagayama R. S. The Uncanny valley: The effect of realism on the impression of artificial human faces. *Presence: teleoperators and virtual environments*. 2007. V. 16. No. 4. pp. 337–351.

32. Steckenfinger S.A., Ghazanfar A.A. Monkey visual behavior falls into the uncanny valley. *Proceedings of the national academy of sciences*. 2009. V. 106. No. 43. pp. 18362–18366.

33. Tinwell A., Grimshaw M., Abdel Nabi D., Williams A. Facial expression of emotion and perception of the uncanny valley in virtual characters. *Computers in human behavior*. 2011. V. 27. No. 2. pp. 741–749.

34. Tinwell A., Grimshaw M., Williams A. The Uncanny wall. *International journal of arts and technology*. 2011. V. 4. No. 3. pp. 326–341.

35. Tinwell A. The Uncanny valley in games and animation. *Boca raton, CRC Press*. 2014.

36. Tinwell A., Sloan R.J.S. Children's perception of uncanny human-like virtual characters. *Computers in human behavior*. 2014. V. 36. pp. 286–296.

37. Tinwell A., Grimshaw M., Abdel Nabi D. The Effect of onset asynchrony in audio visual speech and the uncanny valley in virtual characters. *International journal of mechanisms and robotic systems*. 2015. V. 2. No. 2. pp. 97–110.

38. Thompson J.C., Trafton J.G., McKnight P. The perception of humanness from the movements of synthetic agents. *Perception*. 2011. V. 40. No. 6. pp. 695–704.